Pergamon

Learning and Instruction, Vol. 5, pp. 269–289, 1995

0959–4752/95 $29.00

0959–4752(95)00019–4

KNOWLEDGE DEVELOPMENT AND RESTRUCTURING IN THE DOMAIN OF MEDICINE: THE ROLE OF THEORY AND PRACTICE

H. P. A. BOSHUIZEN, H. G. SCHMIDT, E. J. F. M. CUSTERS and M. W. VAN DE WIEL

University of Limburg, Maastricht, The Netherlands

Abstract

How does the knowledge of the medical student, clerk, intern, and registrar develop and how do formal and informal, classroom and experiential learning contribute to this process? These questions were investigated in two experiments, focusing on knowledge restructuring rather than knowledge acquisition. The experiments showed that practical experience plays an important role in knowledge restructuring. The process was, however, not as continuous as was expected. Notably, advanced students appeared to have considerable knowledge about conditions in patients and their environments that can predispose to disease. However, they rarely applied it in clinical reasoning. Contrary to what was found in expert physicians, advanced students' knowledge about enabling conditions seems not yet to be integrated into their other knowledge about diseases.

Introduction

The development of medical knowledge extends over many, many years. It might be supposed that very young children learn from their parents a kind of folk knowledge about health and illness, prevention and treatment. For instance, they learn about fevers, running noses, colds, and pain. They get to know their parents' reactions to them: if you have a fever, you have to stay in bed or at least inside the house; if your nose is running, they constantly wipe it or tell you to blow your nose; if you have a cold, you have to dress warmly; if you have pain somewhere because you fell, they give you a band aid and a kiss for consolation. Children also gather ideas about the seriousness of different health problems: in some cases you will see a doctor, who will examine you and give you pills or a shot. Later during life, children learn about different classes of diseases: infectious diseases, i.e., diseases that affect the whole class in school, such as the 'flu, chicken pox, measles, mumps, scarlet fever, or rubella; they learn about different types of inflammations, e.g., of the throat or the ear; about traumata, like black eyes, bruises, broken legs or burns, etc. Most of these concepts are learned informally, and to a great

Address for correspondence: Dr H. P. A. Boshuizen, University of Limburg, Department of Educational Research and Development, P.O. Box 616, 6200 MD Maastricht, The Netherlands.

extent from experience. Formal teaching of concepts relating to health and illness in Dutch elementary schools is done in three contexts: hygiene (how to prevent disease), first aid (how to act in case of an accident) and biology (parts and structure of the body). In secondary school, the emphasis is on formal biology and health promotion. In the meantime, experience with diseases, their own or those of other people, also increases. Hence, by the time students enter medical school[1], they have acquired a large amount of experiential, informal, sometimes even folk knowledge about diseases; most of them[2] also have a great deal of formal knowledge about the anatomy and physiology of the body, including cell biology. Medical students build on this knowledge base. It has to be validated, extended and organised in such a way that in the end, it is fit for the task of the physician (i.e., diagnosis of a disease and management and treatment of the patient), which is done through formal teaching and training and practical experience.

Medical education, like most curricula for education in the professions, consists of two parts: a theoretical and a practical part. During the theoretical part, students should acquire the concepts and theories which are supposed to be relevant for professional performance. Some of these concepts and theories are considered "the basics" and are typically taught prior to the applied, clinical subjects. The basics are supposed to be required for understanding the pragmatic relations that are found in everyday practice of the profession. For instance, immunology describes and investigates at the organ and cell level how the body reacts to an invasion by foreign bodies (e.g., viruses and bacteria). It explains, for instance, why on such occasions people are running a temperature or have pain at a specific spot. Typically, it is thought that after having acquired the necessary knowledge, students must practise that knowledge during a clerkship or an internship, or more generally during a period of apprenticeship. Not surprisingly, different kinds of teaching and learning take place during these periods from which different learning outcomes ensue. Formal teaching emphasises the acquisition of formal-domain concepts and theories. These concepts may appear simple to the domain expert. However, for the relative novice in a domain, such concepts can be very hard to learn, and hence hard to teach, because they are often very complex, involving other, related complex concepts (Carroll, 1964). Consequently, students spend much time elaborating on such information and acquiring these concepts. Domain experts who work with these students often cannot appreciate the problems and difficulties of this initial learning and can be puzzled by the banality of questions students ask, by the superficial understanding of the subject matter they are satisfied with, and by the non-discrimination of details of key concepts (Moust, 1993); students do not go beyond the information given. However, even this piecemeal, but uncomplicated, learning can become very complex when new information does not fit the already existing knowledge. This can happen when a student has a misconception about a domain concept or model, e.g., due to the application of analogies which are too simple. Spiro, Feltovich, Coulson and Anderson (1989), for

[1]In the Netherlands, students typically enter medical school at the age of 18 after completion of secondary education at the gymnasium or grammar school level.

[2]Final examinations in grammar school are taken in seven or eight subjects, including Dutch and at least one other modern language as required subjects. Other subjects might be modern languages such as French, German, or even Spanish or Russian, the classic languages Greek and/or Latin, mathematics, economics, history, history of art, geography and geology, physics, chemistry, biology, through to the creative arts and physical education. Universities can require students to have taken exams in a maximum of two other subjects. Medical schools require chemistry and physics, *not* biology.

instance, report misconceptions about the bloodstream in elastic blood vessels resulting from the analogy of the flow of liquids in inflexible pipes. Misconceptions may also stem from deeply entrenched pre-instructional knowledge, such as folk opinions about fever or diarrhoea that are inconsistent with scientific findings (Chinn & Brewer, 1993). These authors also point out that ontological beliefs and epistemological commitments play a role in knowledge revision. Examples of the latter problem are often demonstrated by medical students, who seem to believe that theories must be "true" and they sometimes resist the idea of having to study competing theories. A typical reaction is "Why should we study all those (often sociological or psychological) theories and try to make sense out of these as long as the scientists themselves disagree?" Such an attitude hinders the acquisition of a flexible body of knowledge.

Difficulties like these can arise within one domain, e.g., astronomy (Vosniadou & Brewer, 1987) or muscle physiology (Spiro *et al.*, 1989). However, most curricula require further learning from their students. They have to integrate knowledge from different domains with different paradigms. Existing compartmentalisation must be overcome. For instance, in medicine, students must integrate their knowledge about the production and destruction of cells in general, and of red blood cells in particular, with clinical knowledge about several kinds of anaemia. In order to reach this integration, students must have a good mastery of both (or even more) domains and they must be able to match concepts from one domain with the other, while dealing with loose ends. This is not always possible. Some discrepancies are due to the students' own lack of knowledge; they can be remedied after further study. Other questions are mysteries to the domain specialists as well. For instance, no physician can tell why a patient with anaemia due to vitamin B deficiency can also have neurological problems with his/her feet, but not with his/her hands. This process of adding, integrating and reorganising knowledge is an important enterprise on the way to complete mastery of the domain. According to Rumelhart and Norman (1978), formal learning consists of cycles of knowledge accretion, restructuring and tuning.

Not only do students acquire conceptual knowledge, they also have to learn to use that knowledge in tasks relevant to their future profession. In medicine, they must learn to take a patient's history, do a physical examination, and order lab tests in order to make a diagnosis. These responsibilities require a base of well-organised knowledge, as well as cognitive, perceptual and psychomotor skills. Physicians must also be able to decide on patient management and treatment, and they need the knowledge and skills to communicate to the patient their assessment of the situation and the strategy to be taken. Justification and explanation, adapted to the level of comprehension of the patient, play an important role in patient compliance. Typically, the knowledge and skills needed for these tasks are not learned through direct teaching, nor are they learned from books, but through demonstrations, practical training, and while the students are immersed in medical practice during their clerkships or internships (Patrick, 1992). An important aspect of at least part of this learning is that it occurs in an implicit way and results in tacit knowledge that is hard to verbalise and externalise (Reber, 1989). Jones, Calson and Calson (1986) have shown that acquisition of formal knowledge through clinical experience was very limited.

Not only procedural learning, but also perceptual learning, is important in medicine. A great deal of medical perceptual learning is formalised. Students explicitly learn to differentiate and recognise heart sounds, patterns on ECGs, the structures they feel during physical examination, etc. However, the less obvious, more intricate stimulus configurations, especially those involving more than one sensory system, must be often picked up from the environment. Examples are the patient with diabetes type 1 with poor metabolic control, who is smelling of acetone, or the hypothyroid patient with her or his deep grating voice, slow speech and yellowish skin, or the "secret drinker" with liver problems who has a specific smell that is often masked by perfumes, mouth rinse or after-shave and who has a specific appearance characterised by a yellow shade of the whites of the eyes. Linked to conceptual and procedural knowledge, this perceptual learning plays an important role in the development of diagnostic and patient management proficiency.

These three processes — conceptual, procedural and perceptual learning — play a role in the development of medical expertise. Due to these learning processes, knowledge structures develop and change, finally becoming very flexible and well-versed, easily applicable in routine and non-routine situations. How these changes occur has to be described and explained.

Development of Medical Expertise

When experienced physicians diagnose routine cases, they typically activate their most likely hypotheses in the first minutes of the encounter (Gale & Marsden, 1982). Often they do this on the basis of the patient's appearance and the complaint. Medical, psychological and social information that the doctor knows from earlier contacts with the patient may be used as well (Hobus, Schmidt, Boshuizen, & Patel, 1987). As a next step, the doctor tries to evaluate these hypotheses based on information that is gathered through history taking, physical examination and lab investigation, if necessary. The knowledge applied can be labelled "clinical". If no unexpected information is observed, this procedure can be completed rather quickly. Medical students and inexperienced physicians, on the other hand, take more time for hypothesis generation and information gathering, and they are less certain about their diagnoses (Custers, 1995). Contrary to experts, they apply a fair amount of biomedical and other basic science knowledge (Boshuizen & Schmidt, 1992).

During development of medical expertise, knowledge is accumulated and restructured. In our theory, two ways of restructuring are discerned: knowledge encapsulation and illness script formation. In the next paragraphs, these processes will be described. In the course of their medical training, especially in the first four years of the curriculum, students develop rich, elaborated causal networks explaining the causes and consequences of disease in terms of general underlying biological or pathophysiological processes integrating knowledge from several domains. The result of such network construction is that fourth-year students are able to set up lines of reasoning that can connect such diverse findings as "55-year old man", "pain deep in the abdomen and perineum", "the pain is worse while sitting straight than while lying down", "painful, frequent, difficult micturition" and "soft, enlarged prostrate". In the think-aloud protocol of a fourth-year student[3] who reasoned about such a case, we find such concepts as

[3]These examples are taken from two unpublished protocols of a fourth-year student and an experienced family physician who reasoned about a case of prostatitis in a think-aloud experiment.

prostrate, location of the organs deep in the abdomen like prostrate and urethra, pressure exerted on organs by the gravitational force, enzyme release in infection, and obstruction of the urethra by prostrate swelling. Furthermore, the line of reasoning was not at all focused on the prostrate exclusively, but many other hypotheses concerning over ten other organs were activated during the protocol.

Elaborate application of biomedical knowledge is seldom found in expert protocols (Boshuizen & Schmidt, 1992). For instance, a physician who diagnosed the same case referred to an anatomical structure only once, when he said that he expected that the median lobe of the prostate would be enlarged (and not the right one as could be concluded from the case description). Instead, he used concepts such as urine retention. Furthermore, he was very focused on the prostate as the source of the patient's problems and came to a conclusion in far less time than the student.

These differences between the protocols are a consequence of knowledge restructuring that has resulted from extensive and repeated application of knowledge, restructuring which has occurred particularly through exposure to patient problems.[4] This restructuring which has taken place in the course of the development from novice to expert has led to abbreviations in the lines of reasoning set up in problem solving. The wealth of detailed (patho)physiological and anatomical concepts found in the novice's protocol diminishes, while other types of concepts appear instead, the encapsulating diagnostically relevant concepts that more or less summarize the detailed ones. These encapsulating concepts subsume entire declarative detailed networks of primarily biomedical concepts. Typically, encapsulations are diagnostically relevant concepts of a high level, or simplified causal models explaining signs and symptoms. Encapsulations function as a sort of shorthand in clinical reasoning processes. For example, the term "urine retention" is applied by experienced physicians to explain several symptoms of a patient: it encapsulates the process of a swollen prostate obstructing the urine flow through the urethra which goes through the prostrate, leading to retention of the urine in the bladder, and to anatomical changes in the wall of the bladder, a distended bladder, difficulty with micturition, discomfort, and so on. These encapsulated structures are the results of prolonged knowledge application in the context of diagnostic task execution, hence the diagnostic relevance of encapsulating concepts.

In terms of the learning processes described earlier in this article, knowledge encapsulation is an advanced form of the restructuring phase in a cycle of conceptual learning. By its use, declarative knowledge structures become entrenched in declarative concepts with a higher level of efficiency and practicality. Encapsulating concepts do not necessarily have to be more abstract in the sense of being higher in a hierarchy, although this might be the case. For instance, in our view, the concept "mammal" is not so much an encapsulation because it subsumes all the animals which are members of that class, but rather because it encapsulates the knowledge that there are animals that have a body temperature that is relatively independent of the environmental temperature, that have a (more or less) furry skin, that do not lay eggs but have live-born babies that are nursed with milk produced by a special gland in the mother, etc. In the same vein, "urine retention" is not an encapsulation because it is hierarchically at a higher position than

[4]Patient problems do not necessarily have to be real patients. Life and paper simulations can serve the same function.

the different symptoms it clusters together, but primarily because it summarises a whole pathophysiological process leading to those symptoms.

The reader may have noticed some parallels with Anderson's (1983, 1987) concept of knowledge compilation. We, however, emphasise the conceptual or declarative nature of this process of knowledge change. Contrary to Anderson, we do not assume that the application of declarative knowledge in problem solving generally leads to a transformation into procedural knowledge as a necessary step toward knowledge compilation. We propose that knowledge also can be "compiled" in a declarative format. Encapsulation can be considered an analogue of compilation. For a more detailed discussion of encapsulation, see Schmidt and Boshuizen (1993).

At the same time, and through exposure to patient problems, an extensive set of illness scripts[5] develops. An illness script is a knowledge structure containing a wealth of clinically relevant information about a disease, its consequences (e.g., the complaint a patient brings to the doctor, or the signs and symptoms of a disease during the successive stages of its course) and enabling conditions, the context in which the illness develops (e.g., the physical characteristics of a patient's environment, his or her age, habits, medical history, etc.). These consequences and enabling conditions are linked together with relatively little formal knowledge (compared to what experts have learned about the subject) about pathophysiological causes (the fault) of symptoms and complaints. While solving a problem, a physician searches for an appropriate script and, when he or she has selected one (prostatitis in our example) or a few, he or she will tend to match its elements to the information provided by the patient, hence the focused protocol of the expert. In the course of this script-verification process, the script becomes instantiated. The illness script that is best instantiated will be favoured among the possible diagnoses. Instantiated scripts, in turn, remain available in memory as episodic traces of previously analyzed patients and can be used in the diagnosis of future similar problems.

Illness scripts develop around conceptual clusters of knowledge about diseases (more or less encapsulated, depending on the level of learning) describing the fault at hand. These clusters are linked to features of patients, their environment, family, history, etc., i.e., the consequences and enabling conditions of the disease. They have emerged from continuing exposure to patients and are, therefore, largely the result of extended practice. As a consequence, they reflect the stimulus patterns to which the experts have been exposed in their environment. Although the core of these knowledge clusters has been learnt through conceptual learning, often by formal teaching, a great deal of the learning processes in illness script formation occurs informally, often through perceptual learning. Examples are the smell, sound and physical appearance described earlier in this article. Perceptual learning, however, is not the only factor. Other important processes in illness script development are pattern learning (e.g., the alcoholic patient with the mental, social and physical risks he or she runs) and fine-tuning of the values a specific variable may take in a specific disease. An example of the latter is the lab finding of a serum amylase level of 32 E, which is higher than normal, but is not at all high for the case of acute pancreatitis; it is rather related to chronic pancreatitis. Pattern learning and fine tuning can have both perceptual and conceptual components.

[5]The terms "illness script" and its constituents "enabling conditions" "fault" and "consequences" have been coined by Feltovich and Barrows (1984). Contrary to these authors, we assume that illness scripts are stored and retrieved as such, and are not newly constructed when solving a patient problem.

Illness scripts can only develop as a result of conceptual learning and practical experience. Due to these processes, illness scripts are enriched with experiential knowledge, while the relative amount of formal/causal knowledge decreases. Novices in medicine do not have this kind of illness script, although their lay knowledge may have a similar structure.

Since illness scripts develop around clusters of conceptual knowledge, they are part of a larger conceptual structure. They appear to be mutually linked, based on elements they have in common. For instance, illness scripts may have enabling conditions in common (e.g., the diseases associated with smoking), share complaints (e.g., diseases with a cough), involve the same organ system (respiratory diseases), have symptoms in common (e.g., cyanosis), or may be based on the same pathophysiology (e.g., diseases with left ventricular inflow obstruction). Even very pragmatic links can be found, as is suggested by statements such as "In case of an adult with thrush[6] you should always think of an underlying disease like diabetes, or the patient might have recently undergone some treatment with antibiotics." This link structure is not at all static. Practical experience can result in addition to and change in these relations (see Hassebrock & Prietula, 1992).

Illness scripts seem to exist at various levels of generality, ranging from representations of general disease categories such as respiratory disease to representations of individual patients seen before, such as the four-year old boy with severe shortness of breath and fever with no history of chronic respiratory disease.

A final aspect of our theory is the assumption that the knowledge structures acquired during the different stages of development — pathophysiological networks, encapsulated structures, illness scripts and instance scripts — do not decay, nor do they become inert, or inaccessible. They sediment into multiple layers which are accessed when ontogenetically more recently acquired structures fail in producing an adequate representation of a clinical problem (Schmidt, Norman, & Boshuizen, 1990).

The central concepts in this theory are knowledge encapsulation and illness script formation. Two experiments will be described that provide empirical evidence for these processes.

Experiment 1

The phenomenon of knowledge encapsulation was first investigated by Schmidt and Boshuizen (1993). The general idea was that the initial stages of conceptual learning in medical school are characterized by knowledge accretion, leading to an increase in the number of concepts that can be applied in medical problem solving. A further increase in expertise is, however, associated with encapsulation of concepts in the knowledge base, resulting in the subsuming of a great number of detailed biomedical concepts under higher level, clinically relevant concepts. Hence, at the lowest levels of expertise, where learning is best characterized by knowledge accretion and when no encapsulation has yet taken place, reasoning will show an increase in the number of concepts applied with increasing expertise. However, an expert's reasoning will contain less detailed biomedical concepts and more encapsulating concepts than the reasoning of a less experienced

[6]Thrush is an infection in the mouth especially found in infants.

subject at an intermediate level of expertise. This sequence of knowledge accretion and encapsulation would hence result in an inverted U-shaped relation between medical expertise and the number of concepts applied in the pathophysiological explanation of the signs and symptoms in a medical case. This relation was indeed found. The increase in the number of encapsulating concepts was also found. The number of detailed concepts was not investigated.

Furthermore, Schmidt and Boshuizen hypothesised that, since experts activate whole illness scripts, the relevant encapsulated pathophysiological concepts are also activated at once. Students, on the other hand, have to reason through their biomedical knowledge, in order to build a model that can explain the patient's signs and symptoms. The concepts applied will be activated one by one. This model building takes time, much more time than illness script activation and instantiation. As a consequence, reduction of case reading time would seriously affect the student's explanation, while the expert's performance would hardly suffer. This relation was indeed reported: diagnostic accuracy increased with increasing levels of expertise. Furthermore, experts' diagnostic performance and their explanations were not affected by a sizable reduction in processing time, whereas the students' explanations were. However, in this study only one case had been used, a case of a disease that had shown remarkable changes in incidence rate over the years it had been used in research, both by us and by Patel and Groen (1986). Since the experiment plays a pivotal role in the confirmation of our theory, there was a need to replicate it with different cases.

Method

This experiment was replicated with two cases, taken from two different specialty areas in internal medicine, four levels of expertise, and three reading time conditions. The subjects were 96 students and physicians at the University of Limburg: 24 second-year, 24 fourth-year and 24 sixth-year medical students and 24 internists with at least four years of experience in internal medicine.[7] Each group of 24 was randomly subdivided into three groups of eight which were assigned to three time constraint conditions. Subjects received a small compensation for their participation.

The materials consisted of two booklets, each containing a description of a clinical case and two blank response sheets. Each clinical case description reported some contextual information and the complaint, the case history, the physical examination, the relevant laboratory data and some additional findings. The case descriptions were about half a page in length. The two clinical cases, pheochromocytoma and liver cirrhosis, were based on actual patients and were presented as a narrative. An example of such a case description can be found in Appendix 1.

Subjects were asked to produce a diagnosis and explain the pathophysiological process underlying the case. Depending on the experimental condition, subjects were given the opportunity to read each case for 3 minutes (3'00"), 1 minute and 15 seconds (1'15"),

[7] The original experiment which is replicated in this study also included a group of allied health sciences students, providing a sort of base line.

or 30 seconds (30″). Subjects were free to use as much time as they needed for the assignments. The order of case presentation was balanced.

Diagnoses were scored on a scale ranging from 0 (incorrect diagnosis) to 6 (completely correct diagnosis) for each case. The pathophysiological explanations were schematised as semantic networks. The total number of concepts in these schematised explanations was counted. For each case, an ideal explanation was constructed (see Appendix 2). This explanation contained the minimal set of pathophysiologically and clinically relevant concepts necessary to explain the signs and symptoms in the case. The subjects' explanations were matched against the ideal explanations: the number of concepts in the explanation which were identical or equivalent to concepts in the ideal explanation was counted (termed "model concepts"), as was the number of explanation concepts that were at a more detailed level. These three measures, in isolation and in combination, reflect knowledge encapsulation. The number of concepts in the pathophysiological explanations will first increase and subsequently decrease. The number of model concepts will increase while the detailed concepts should decrease. A worked out example of this scoring procedure can be found in Appendices C–D. The data were analyzed with a 3-factor ANOVA with repeated measurements over one factor (case).

Results and Discussion

Diagnosis

Figure 1 shows the diagnostic accuracy of the different groups. As in the original experiment by Schmidt and Boshuizen (1993), diagnostic accuracy increased with increasing levels of expertise ($F(3, 84)=37.151$, $p=0.0001$). Again, reading time constraints had no significant effects on this measure ($F(2, 84)=1.560$, $p=0.216$).

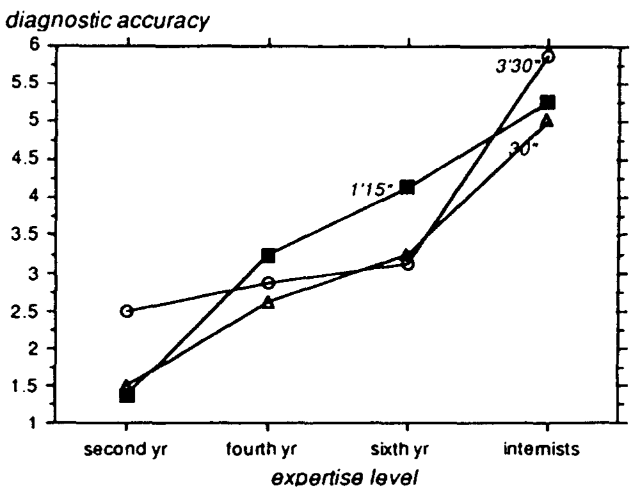

Figure 1. Accuracy of the diagnoses given by subjects of four different levels of medical expertise of a case studied under three different time conditions.

Case Explanation

As shown by Figures 2–4, case explanation varied with level of expertise. Figure 2 shows the number of concepts used in the pathophysiological explanations. Like in the original study, the total number of concepts had an inverted U-shaped relationship with level of expertise ($F(3, 84)=8.61$, $p=0.0001$). Furthermore, Figure 3 reveals that the number of model concepts applied in the case explanations increased monotonically with expertise level ($F(3, 84)=19.37$, $p=0.0001$), while the number of more detailed biomedical concepts applied had an inverted U-shaped relation with expertise level ($F(3, 84)=3.87$, $p=0.012$): advanced students used more detailed concepts than experts and junior students (Figure 4). These three findings confirm the knowledge encapsulation hypothesis: after a stage of knowledge accretion, further development is characterized by the encapsulation of biomedical knowledge, leading to an increase in higher level concepts and a decrease in detailed biomedical concepts in clinical reasoning.

The results of the constraint on reading time were, however, not so clear. Subjects at lower and intermediate levels of expertise development would be susceptible to this manipulation, but at a more advanced stage, knowledge encapsulation and illness script formation would lead to a relative insensitivity. As expected, there was an effect of reading time constraint on the number of concepts used: the more time available for studying the cases, the more extensive the explanation was ($F(2, 84)=3.623$; $p=0.031$). However, Figure 2 shows that this was irrespective of expertise level: no interaction effect was found ($F(6, 84)=1.020$; $p=0.4179$). In the original Schmidt and Boshuizen (1993) study, the extent of the explanations generated by the second- and fourth-year students was heavily affected by time constraint, while the explanations of the other groups were unaffected. Comparisons per expertise level in the present experiment revealed that none of the simple effects showed a significant difference. The group of fourth-year students came closest ($F(2, 21)=2.638$; $p=0.0950$), but all other p-values

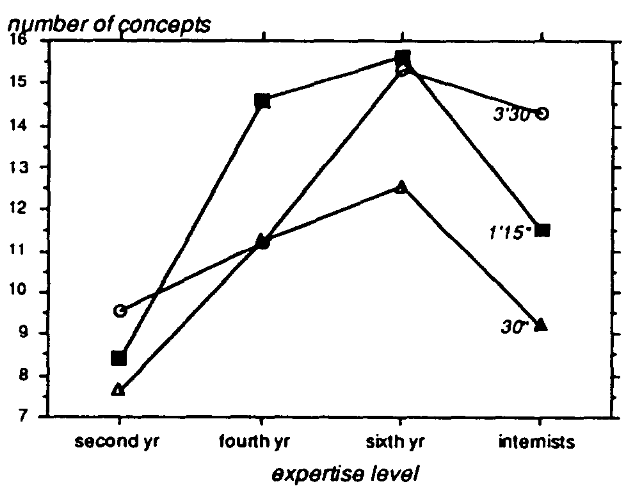

Figure 2. Elaborateness of the explanation given by subjects of four different levels of medical expertise of a case studied under three different time conditions.

were >0.10. This might at best indicate a trend in the direction expected, however, not a very strong one.

Figure 3 shows that the amount of processing time allowed also had a significant effect on the number of model concepts ($F(2, 84)=9.77$, $p=0.0001$), which was also the case in the original experiment. No interaction effect was found ($F(6, 84)=1.139$; $p=0.2577$). Planned comparisons per expertise level were all significant ($p<0.05$), except for the expert group ($F(2, 21)=2.473$, $p=0.1085$). Time constraint had no effect on the number of detailed concepts applied (main: $F(2, 84)=0.300$; $p=0.7414$; interaction: $F(6, 84)=0.254$; $p=0.9564$) (see Figure 4).

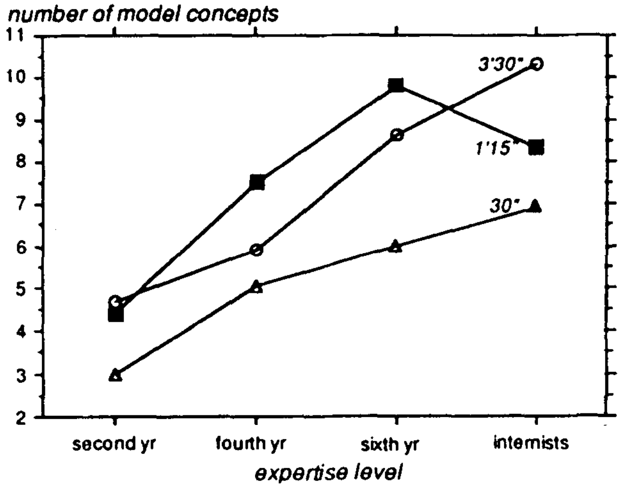

Figure 3. Number of model concepts in the explanation given by subjects of four different levels of medical expertise of a case studied under three different time conditions.

This combination of results complicates the assessment of the confirmation level of the knowledge encapsulation hypothesis. The overall effects of level of expertise on the number of concepts, number of model concepts, and number of detailed concepts applied unambiguously confirm the hypothesis. However, the effects of reading time manipulations are not so clear, as far as the number of concepts applied in the explanations are concerned; they show only a trend in the direction expected. The number of model concepts replicates the original findings. So far, it is difficult to decide on the importance of these outcomes. Not only is the theory rather new, but so are the procedures applied to investigate it. No further research is available with which to compare the present outcomes. The outcomes of the present and original experiments are summarised in Table 1.

Differences in difficulty between the two cases in this experiment and also in comparison to the endocarditis case originally used, may have played a role in these outcomes. The liver cirrhosis case seemed easiest. Diagnosis scores were better for the liver cirrhosis case than for the pheochromocytoma case ($F(1, 84)=39.116$; $p=0.0001$), though this was not the same with the experts (interaction $F(3, 84)=8.450$; $p=0.0001$). The endocarditis case may have been even more difficult, especially for the second- and

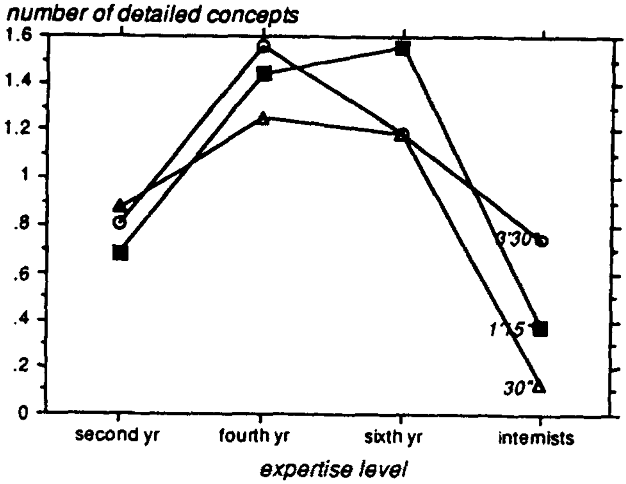

Figure 4. Number of detailed, biomedical concepts in the explanation given by subjects of four different levels of medical expertise of a case studied under three different time conditions.

Table 1
Comparison of the results of Experiment 1 and the outcomes by Schmidt and Boshuizen (1993)

	Dependent variable used		
Experiment	No. of concepts	No. of model concepts	No. of detailed concepts
Schmidt and Boshuizen (1993)			
Expertise	Inverted U-shaped	Linear increase	Not investigated
Time	Yes	Yes	Not investigated
Interaction	Yes	No	Not investigated
Experiment 1			
Expertise	Inverted U-shaped	Linear increase	Inverted U-shaped
Time	Yes	Yes	No
Interaction	No	No	No

fourth-year medical students; e.g., no second-year students and only a few fourth-year students had recognised at least one important aspect of the diagnosis, but students in the present study could do that. This phenomenon could explain why students are not so susceptible to time constraint manipulations as in the original experiment, but it does not clarify why experts also seem only slightly affected by differences in reading time.

Experiment 2

In the introduction of this paper it was conjectured that illness scripts develop around conceptual clusters of knowledge about diseases (more or less encapsulated depending on the level of learning) describing the fault at hand, and that illness script development is primarily affected by practical experiences with patients in a diagnostic and/or therapeutic setting. We have also assumed that illness script formation and knowledge encapsulation are simultaneous processes. Probably, the driving force for encapsulation is active application of causal knowledge in diagnostic and therapeutic problems, while experience with real patients feeds illness script formation, enrichment and tuning. During the first years of a medical curriculum, such practical experiences are rare, although the newer approaches to medical teaching use realistic simulations as soon and as often as possible. Hence, the development of rich, experientially grounded illness scripts cannot be expected before the clerkships. In particular, this will be the case for the enabling conditions of the illness script which receive less attention in the handbooks than the consequences of a disease. Experiment 2 addresses the hypothesis that experts have illness scripts that are superior regarding enabling conditions.[8]

Method

The method applied was derived from Hobus, Boshuizen and Schmidt (1990), who investigated the structure of illness scripts by asking subjects to describe prototypical patients with a specific disease (cf. Abelson, 1981). Description of a prototypical patient entails, however, a methodological problem. It might induce the subjects to produce relatively more patient description statements, thus affecting less experienced subjects in a different way than experts who have seen more patients. Therefore, a more neutral probe was also applied, i.e., half of the subjects were asked to describe the clinical picture of a specific disease.

The subjects were 22 fourth-year students, 22 sixth-year students, 23 people who had recently graduated from medical school and were training to be family physicians (interns), and 22 experienced family physicians. Subjects were asked either to describe prototypical patients with a specific disease, or to describe the clinical pictures of those diseases. The disease names were taken from the cases used in an earlier experiment.

The descriptions were divided into separate idea units that were classified as either enabling conditions (when referring to features and circumstances of the patient, and his or her medical and non-medical history), faults (when referring to pathophysiological processes associated with the disease) or consequences (when referring to signs and symptoms displayed). About 6.5% of the idea units, predominantly referring to patient management and treatment, could not be assigned to any of these classes.

Results and Discussion

Enabling Conditions

Figure 5 describes the proportion of enabling conditions in the descriptions of prototypical patients or clinical pictures given by subjects of different levels of expertise. Analysis revealed that the proportion of idea units reported referring to enabling conditions increased with increasing levels of expertise ($F(3, 81)=9.152$; $p<0.001$),

[8]These data have been presented earlier at the Annual Meeting of the American Educational Research Association, San Francisco, CA, 20–24 April 1992.

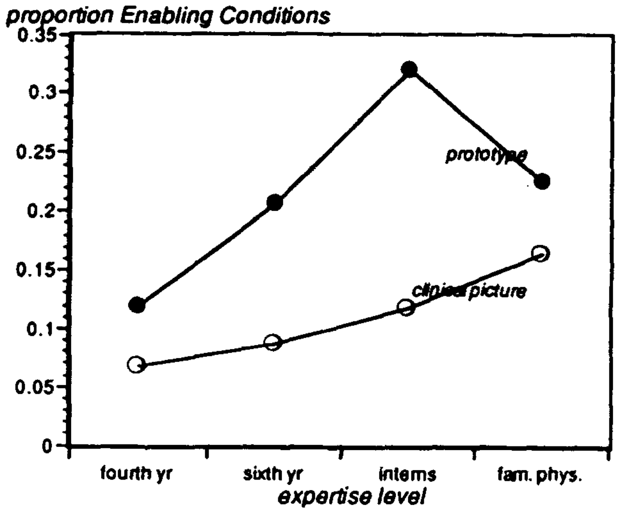

Figure 5. Proportion of enabling conditions idea units in the descriptions of prototypical patients and clinical pictures.

suggesting illness script enrichment. The kind of probe also affected the proportion of enabling conditions idea units mentioned (F (1, 81)=37.692, p<0.001). Subjects who described prototypical patients mentioned more idea units than subjects describing clinical pictures. Furthermore, an interaction between probe and expertise level was found (F (3, 81)=3.549; p<0.05): especially for intermediates who produced far more enabling conditions when asked to describe a prototypical patient than when asked to describe the clinical picture of the disease. This suggests that the intermediates have this kind of knowledge, but that it is not yet fully integrated into their illness scripts.

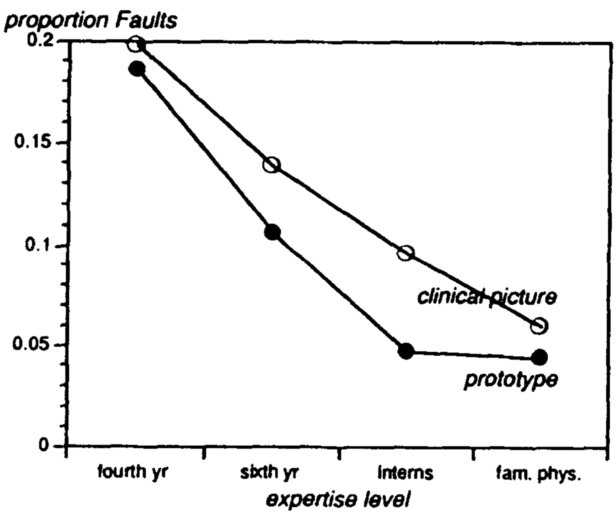

Figure 6. Proportion of fault idea units in the descriptions of prototypical patients and clinical pictures.

Fault

Figure 6 shows the proportion of fault statements that decreased monotonically ($F(3, 81)=16.137$; $p<0.001$) with increasing expertise levels. No effects (neither main nor interaction) of the kind of probe applied could be found. Regardless of whether subjects described the clinical picture of a disease or the prototypical patient having that disease, the proportion of Fault statements remained the same in their narratives.

Consequences

Differences between the proportions of consequences reported by the different expertise groups are surprisingly small (Figure 7). The means vary only between 65.1% for the interns and 67% for the sixth-year students. However, it was again found that the kind of probe applied affected the relative number of consequences mentioned ($F(1, 81)=4.764$, $p<0.05$). Subjects who described clinical pictures reported relatively more consequences than the others. The interaction between expertise and kind of probe was significant ($F(3, 81)=2.711$; $p<0.05$); physicians and fourth-year students were not affected by probe differences, while sixth-year students and interns were.

These results are remarkable in more than one way. First, the steady decrease of fault elements in the descriptions, not affected by the kind of probe applied, is a further confirmation of the hypothesis of knowledge encapsulation. This more or less encapsulated biomedical knowledge connects the other elements in the illness script. An example, taken from the description of a clinical picture of Herpes Zoster given by a fourth-year student, can illustrate this:

> It's the chicken pox virus, you've had chicken pox in your youth, and the virus settles down in your nerves, and you get vesicles on your skin, and these are very contagious, because the fluid contains the virus, and you'll never get rid of it; and it can cause an infection of the eyes, and that might be very dangerous.

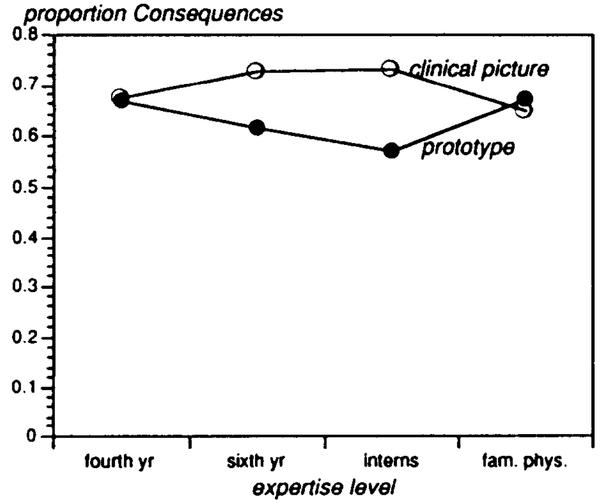

Figure 7. Proportion of consequences idea units in the descriptions of prototypical patients and clinical pictures.

Experts do not refer to the fault in such an elaborate way, but use concepts such as "virus", "(skin)nerve(segment)", and "weak constitution" that serve the same purpose.

The results also show that subjects who are requested to describe a clinical picture are more apt to volunteer relatively (and absolutely) more consequences, while subjects who describe prototypical patients generate more enabling conditions. This was especially the case for the interns and — to a lesser extent — for sixth-year students. Physicians and fourth-year students were hardly affected by the differences in probes applied. This finding was not expected. Remember that we added a second probe, because the family physicians' longer experience might be reflected in a relatively and absolutely higher number of enabling conditions in the prototype descriptions. Apparently, the interns who are training as family physicians have already gathered a great deal of information about characteristics of patients, their environments and/or medical and psychosocial backgrounds. However, this information is only brought forward when they are quite explicitly asked for it, indicating that it is not yet part of an illness script that is activated as a whole. On the other hand, information retrieval in the students' and medical experts' knowledge bases is obviously quite insensitive to the kind of probe applied. This might indicate that they search only one knowledge base irrespective of the question asked: the students search their causal networks, the family physicians their illness scripts. The other groups might search one of at least two, not fully integrated, knowledge structures dependent on the exact wording of the question.

General Discussion

Experiment 1 largely replicated the results found earlier with another case (Schmidt & Boshuizen, 1993). Deviations occurred especially in the effects of reading time constraint on the total number of concepts applied in the case explanation. While in the original study, second- and fourth-year students were extremely susceptible to this manipulation and experts were definitely not, in Experiment 1, only the fourth-year students showed a slight tendency in this direction. Deviations from earlier results might be attributed to discrepancies in case difficulty for the different expertise groups. The diseases used to investigate knowledge encapsulation in Experiment 1 and in the original study are studied at different moments in the curriculum by the students and showed different profiles of diagnosis scores, number of model concepts and number of detailed concepts applied. This suggests that knowledge encapsulation is, in the first place, a result of repeated knowledge application in the context of a diagnostic task — of practical experience. Having seen one or a few cases is not sufficient. This conclusion was further confirmed by Experiment 2.

The role of practical experience in illness script formation is even more conspicuous. Experiment 2 showed an increase in the number of enabling conditions mentioned corresponding with the number of patients students had seen and with the number of years of experience. Seeing patients in a health care setting results in the acquisition of knowledge about the characteristics of those who are especially susceptible for specific diseases. Hobus *et al.* (1987) have shown that this kind of knowledge is a key instrument in the activation of an appropriate set of hypotheses about what a specific patient might have. These authors also found that recent graduates from medical schools are not able to benefit from this kind of information, when it is available. Experiment 2 suggests that this inability does not result from a lack of knowledge. On the contrary, when sixth-year students and interns are asked to "describe as detailed as possible the prototypical patient having . . . (Herpes Zoster, for instance)", they are in no way inferior to the experienced family physicians (either in the number of enabling conditions mentioned, or in the quality). Yet, when asked to "describe as detailed as possible the clinical picture of Herpes Zoster", their attention shifts toward the consequences of that disorder. Apparently, the enabling conditions knowledge is quickly acquired, but the integration into other knowledge leading to stable illness scripts that can be easily activated, even in information-limited situations very early in a doctor patient encounter, requires more and longer experience. Accessibility of illness scripts is not a major issue in our theory, although some research has been done on it (see Hobus, 1994), revealing that physicians learn to use some kinds of enabling condition earlier than others. For instance, the role of sex, age, smoking and alcohol abuse is easily acknowledged, while drug therapy and previous operations, especially in combination with other enabling conditions, are learned later in a career as important factors. The outcomes reported here suggest that more attention to this question might be fruitful.

Acknowledgement — The authors wish to thank Prof. P. W. Leeuw, Dr P. P. M. Hobus and Dr P. A. Hanson. Their help with the construction of the ideal explanation of the liver cirrhosis case and with polishing the English was invaluable.

References

Abelson, R. P. (1981). Psychological status of the script concept. *American Psychologist*, **36**, 715–729.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J. R. (1987). Skill acquisition: compilation of weak-method problem solutions. *Psychological Review*, **94**, 192–210.

Boshuizen, H. P. A., & Schmidt, H. G. (1992). The role of biomedical knowledge in clinical reasoning by experts, intermediates and novices. *Cognitive Science*, **16**, 153–184.

Carroll, J. B. (1964). *Language and thought*. Englewood Cliffs, NJ: Prentice-Hall.

Chinn, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: a theoretical framework and implications for science instruction. *Review of Educational Research*, **63**, 1–49.

Custers, E. J. F. M. (1995). *The development and function of illness scripts; studies on the structure of medical diagnostic knowledge*. Maastricht: Datawyse

Feltovich, P. J., & Barrows, H. S. (1984). Issues of generality in medical problem solving. In H. G. Schmidt & M. L. De Volder (Eds.), *Tutorials in problem-based learning, a new direction in teaching the health professions* (pp. 128–142). Assen: Van Gorcum.

Gale, J., & Marsden, P. (1982). Clinical problem solving: the beginning of the process. *Medical Education*, **16**, 22–26.

Hassebrock, F., & Prietula, M. J. (1992). A protocol-based coding scheme for the analysis of medical reasoning. *Journal of Man–machine Studies*, **37**, 613–652.

Hobus, P. P. M. (1994). *Expertise van huisartsen; praktijkervaring, kennis en diagnostische hypothesevorming*. [*Expertise of general practitioners: practical experience, knowledge and generation of diagnostic hypotheses*]. Amsterdam: Thesis Publ.

Hobus, P. P. M., Boshuizen, H. P. A., & Schmidt, H. G. (1990). Expert–novice differences in the mental representations of patient prototypes. Paper presented at the 2nd International Symposium on Problem-based Learning. Yogyakarta, Indonesia.

Hobus, P. P. M., Schmidt, H. G., Boshuizen, H. P. A., & Patel, V. L. (1987). Contextual factors in the activation of first diagnostic hypotheses: expert–novice differences. *Medical Education*, **21**, 471–476.

Jones, J. G., Calson, G. J., & Calson, C. (1986). The acquisition of cognitive knowledge through clinical experiences. *Medical Education*, **20**, 10–13.

Moust, J. H. C. (1993). *De rol van tutoren in probleemgestuurd onderwijs; Contrasten tussen student- en docent-tutoren*. [*The tutor role in problem-based education; Contrasts between peer and faculty tutors*]. Maastricht: University of Limburg.

Patel, V. L., & Groen, G. J. (1986). Knowledge-based solution strategies in medical reasoning. *Cognitive Science*, **10**, 91–110.
Patrick, J. (1992). *Training: research and practice*. London: Academic Press.
Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, **118**, 219–135.
Rumelhart, D. E., & Norman, D. A. (1978). Accretion, tuning and restructuring: three modes of learning. In R. Klatzky, & J. W. Cotton (Eds.), *Semantic factors in cognition*, (pp. 37–52). Hillsdale, NJ: Erlbaum.
Schmidt, H. G., & Boshuizen, H. P. A. (1993). On the origin of intermediate effects in clinical case recall. *Memory and Cognition*, **21**, 338–351.
Schmidt, H. G., Norman, G. R., & Boshuizen, H. P. A. (1990). A cognitive perspective on medical expertise: theory and implications. *Academic Medicine*, **65**, 611–621.
Spiro, R. J., Feltovich, P. J., Coulson, R. L., & Anderson, D. K. (1989). Multiple analogies for complex concepts: antidotes for analogy-induced misconceptions in advanced knowledge acquisition. In: S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning*, (pp. 498–531). Cambridge: Cambridge University Press.
Vosniadou, S., & Brewer, W. F. (1987). Theories of knowledge restructuring in development. *Review of Educational Research*, **57**, 51–67.

Appendix 1

Liver Cirrhosis Case Applied in Experiment 1

A 45-year old lawyer complains about a nagging pain in the upper abdomen he has been suffering from for some months. He attributes this pain to stress due to a decline in his clientele and his divorce two years ago. He pays occasional visits to a prostitute, but the last time, he appeared impotent. He admits substantial alcohol consumption and smoking a lot (40 cigarettes/day). There is no evident food intolerance.

Physical examination revealed a somewhat extended abdomen, but no shifting dullness. The liver can be palpated, and has an irregular surface; the spleen cannot be palpated. At the ankles a little edema is found. The testes are rather small. The thorax shows a spider naevus.

Laboratory tests showed a ESR of 44 mm/h (normal: <12 mm/h) and a hemoglobin level of 8.0 mmol/l (normal: 8.5–11.0 mmol/l). Sodium 138 mmol/l (normal: 132–142 mmol/l), potassium 3.6 mmol/l (normal: 3.6–5.0 mmol/l), ALT (SGPT) 120 U/l (normal: <40), AST (SGOT) 84 U/l (normal: <40 U/l), LDH (LD) 800 U/l (normal: 200–450 U/l), GTT 250 U/l (normal: <50 U/l), alkaline phosphatase 200 U/l (normal: 30–125 U/l) and bilirubin 25 μmewl/l (normal: <17 μmewl/l).

Appendix 2

The ideal explanation is represented as a network; rectangular frames contain explanatory concepts, frames with rounded corners represent case elements. Most links between concepts are directed (indicated by an arrow), referring to causal, conditional and attribute relations. Identity relations are non-directed. Some links are labelled "cond/neg", indicating that this specific finding contradicts the other concept. Arrows within a frame refer to increase or decrease.

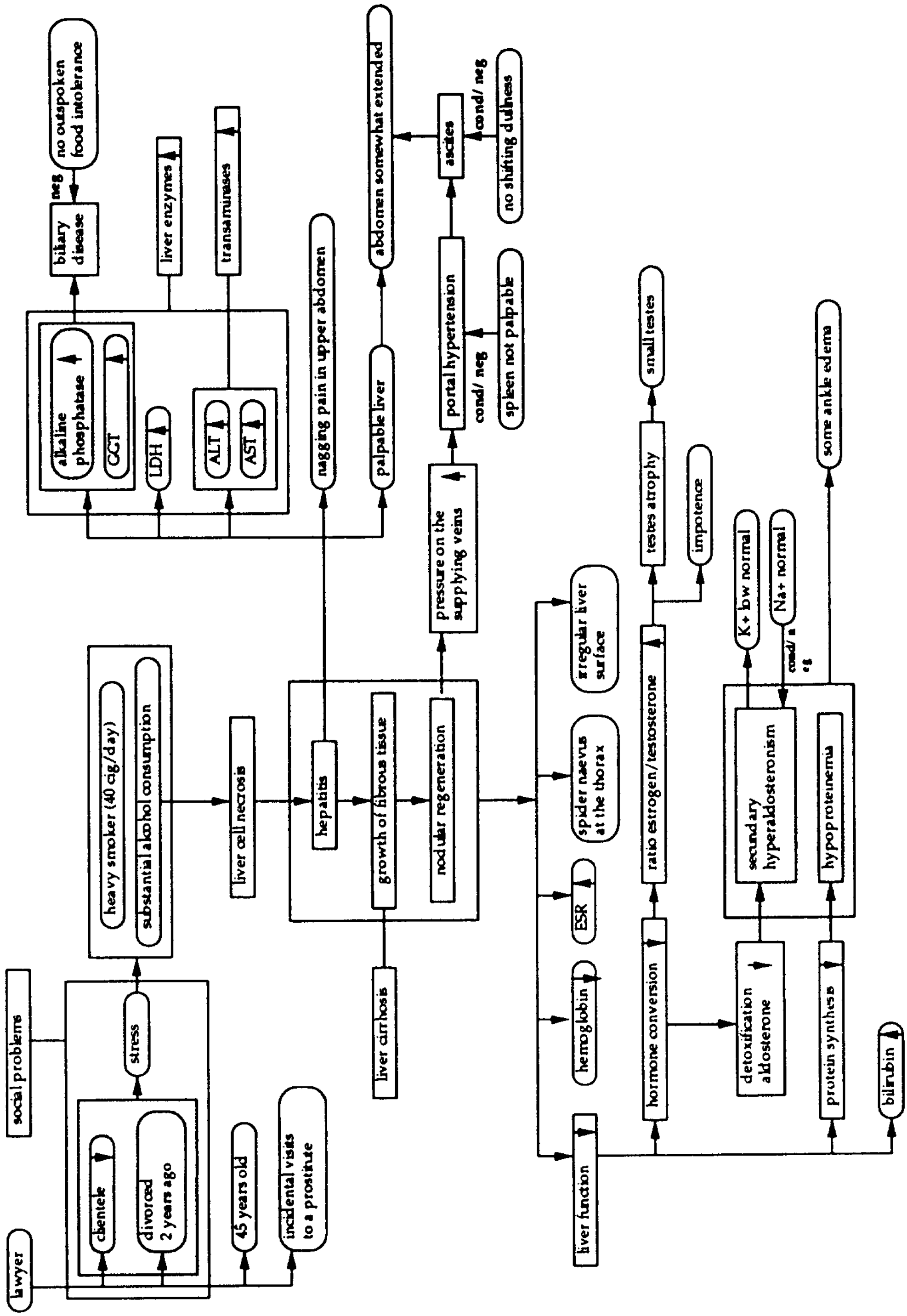

Figure 8. Ideal explanation of the liver cirrhosis case.

Appendix 3

Scoring of an explanation of the liver cirrhosis case

Student Explanation

Due to stress he smokes and drinks a lot. This has resulted in a fatty liver, and hence pressure on the supplying veins. Then blood cannot stream fast through the liver and gets stuck before the liver, causing portal hypertension. This results in excretion of liver fluids from the lymph capsula, leading to ascites and an extended abdomen.

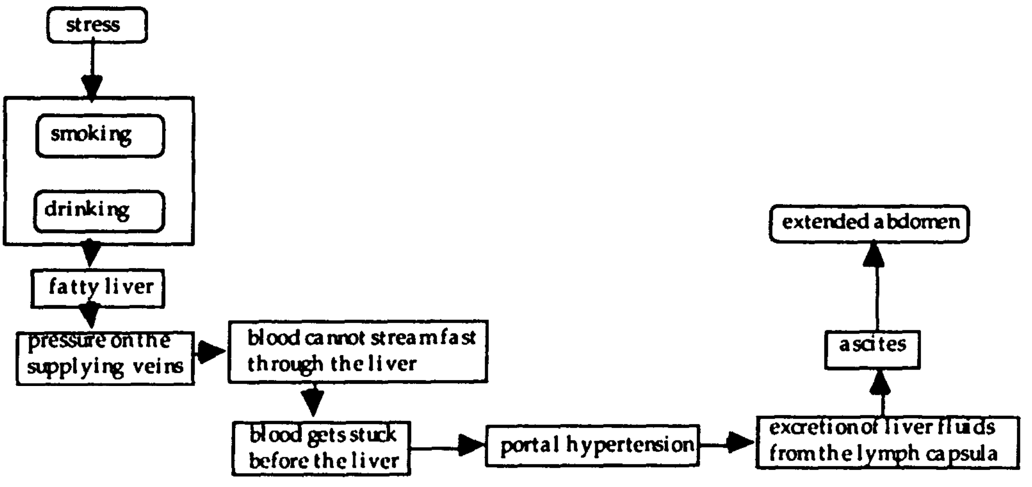

Figure 9. Semantic network derived from the student explanation.

Frame and link codes are similar to those in the ideal explanation, Appendix 2

Comparison with the Ideal Explanation

This semantic network has been projected on the ideal explanation (see Appendix 4). Identical concepts are stress, smoking, drinking (a lot), pressure on the supplying veins, portal hypertension, ascites and extended abdomen. This results in six model concepts.

Fatty liver, blood cannot stream fast through the liver, blood gets stuck before the liver, excretion of liver fluids from the lymph capsula are more detailed concepts used to explain the relation between some of the model concepts. This results in four detailed concepts.

Appendix 4

Identical and equivalent concepts and links are shaded, different concepts and links have been added.

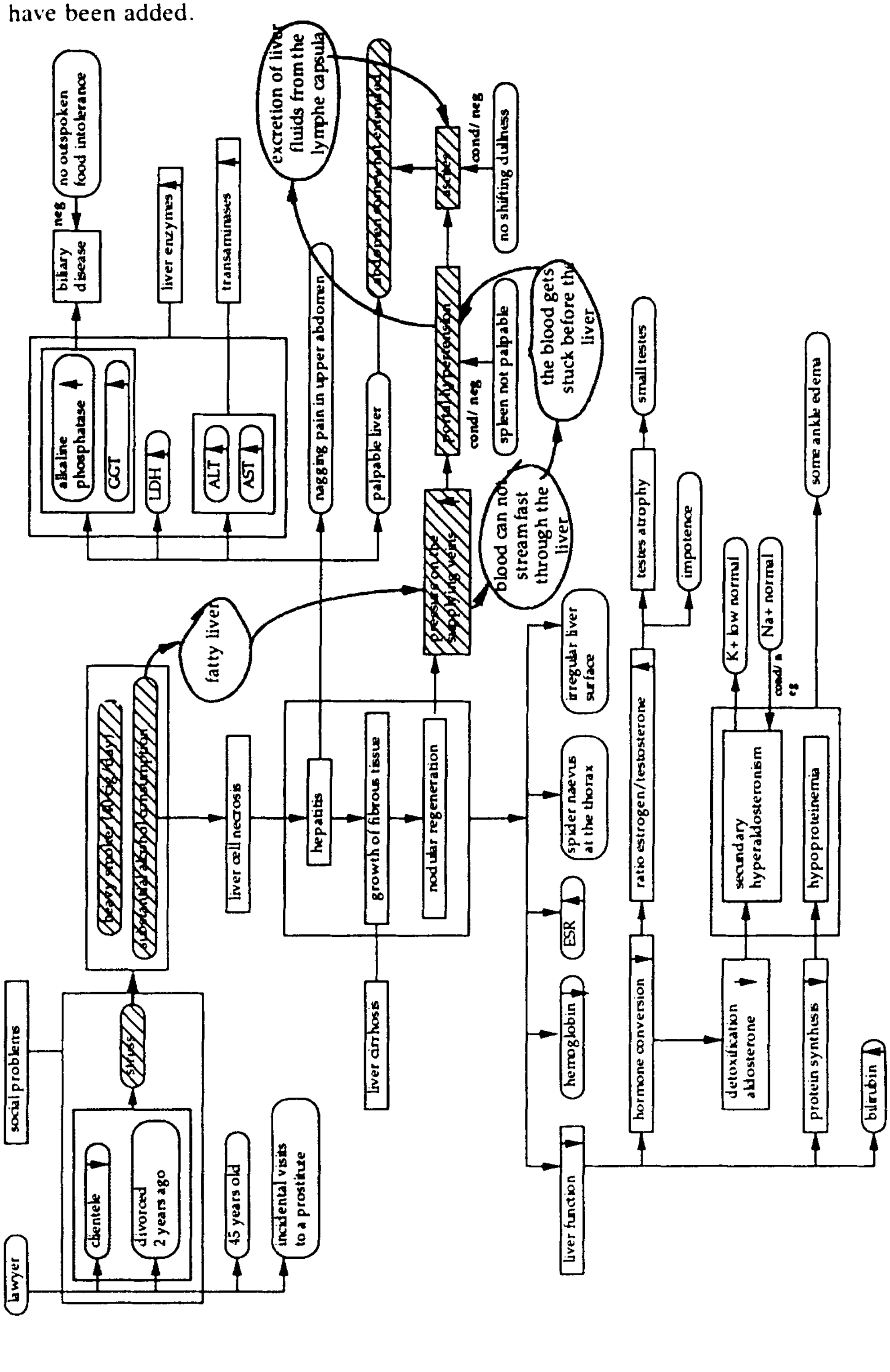

Figure 10. Ideal explanation of the liver cirrhosis case with the student's explanation noted on it.