COGNITIVE SCIENCE 16, 153-184 (1992)

# On the Role of Biomedical Knowledge in Clinical Reasoning by Experts, Intermediates and Novices

HENNY P.A. BOSHUIZEN AND HENK G. SCHMIDT
*University of Limburg, The Netherlands*

In two studies the role of biomedical knowledge in the diagnosis of clinical cases was explored. Experiment 1 demonstrated a decrease in the use of biomedical knowledge with increasing expertise. This result appeared to be at variance with some findings reported in the literature (e.g., Lesgold, 1984), but supported those of others (e.g., Patel, Evans, & Groen, 1989). In Experiment 2, three possible explanations for this phenomenon were investigated: (1) rudimentation of biomedical knowledge, (2) inertia, and (3) encapsulation of biomedical knowledge under higher order concepts. Using a combined think-aloud and post-hoc explanation methodology, it was shown that experts have more in-depth biomedical knowledge than novices and subjects at intermediate levels of expertise. The findings generally support a three-stage model of expertise development in medicine consisting of acquisition of biomedical knowledge, practical experience, and integration of theoretical and experiental knowledge resulting in knowledge encapsulation.

## INTRODUCTION

As early as the 15th century, physicians and other students of human biology tried to peer into the "black box" of the human body. People such as Antonio Benivienis (1448-1502) and Jean Fernel (1506-1588) attempted to relate their clinical observations to pathological-automatic findings obtained from postmortem examinations. Eustachius (1524-1574), Fallopius (1523-1574), and Fabricius ab Aquapendente (1547-1619) described many organs

We wish to acknowledge Geoffrey R. Norman and Pie P.M. Hobus and four anonymous reviewers for their comments on earlier versions of this article. The studies described in this project were enabled by a grant (#6626) from the Institute for Educational Research in the Netherlands (SVO). Preparation of the article was made possible by a Spencer fellowship from the National Academy of Education (USA) to Henny P.A. Boshuizen.

Correspondence and requests for reprints should be sent to Henny P.A. Boshuizen, Rijksuniversiteit Limburg, Educational Development and Research, P.O. Box 616, 6200 MD Maastricht, The Netherlands.

and other structures in the human body, whereas Harvey (1578–1657) discovered blood circulation. Through the efforts of these investigators, the secrets that were kept safe in the "box" were discovered. It was not easy to carry out investigations in those days. Tools were not as sophisticated as they are now. Corpses were scarce, as only criminals sentenced to death were allowed to be dissected. Only in winter could a complete anatomic investigation of a corpse be carried out. However, since the days of Boerhaave (1668–1738), the *biomedical sciences,* such as anatomy and physiology, have been considered crucial to the clinical sciences. Research into the structure and functioning of the human body has provided increasing insight into its normal functioning and maintenance of homeostasis. As a result of these research efforts, the mechanisms that underlie long-known empirical rules of thumb became understood, and medicine developed from an art into a modern science. In particular, since the beginning of this century, the biomedical sciences have played an increasingly important role in the medical curriculum.

Notwithstanding its importance for medicine as a science, the role of biomedical knowledge in medical diagnosis and treatment in everyday practice is not at all clear. Research findings are sometimes contradictory; a theory explaining these diverging research outcomes is still lacking. The aim of the studies to be presented here is to investigate the role of biomedical knowledge in clinical reasoning at different levels of medical expertise, in an attempt to reconcile opposing views. First, however, the role of knowledge in medical diagnosis will be discussed and studies addressing the issue of biomedical knowledge and clinical reasoning will be reviewed.

**The Role of Biomedical Knowledge in Medical Diagnosis**

Cognitive theories of medical diagnosis assume it to be a process of comprehension or problem solving. "Mental representation of the problem" (or "problem representation") is a key concept in these theories. This position has been elaborated by Feltovich and Barrows (1984) who defined a problem representation as "...a cognitive structure corresponding to a problem (which is) constructed by a solver on the basis of domain related knowledge and its organization." The problem representation is "the solver's internal model of the problem, containing the solver's conceptions of problem elements, their relationships to each other, the goals of problem solving, etc." (Feltovich & Barrows, 1984, p. 132; see also Chi, Feltovich, & Glaser, 1981). According to these authors, the representation of a diagnostic problem takes the form of an (instantiated) "illness script" in which patient characteristics, signs, and symptoms of the disease, and knowledge of underlying processes are organized. These illness scripts describe the patient's present condition and how it came to be, understood through biomedical knowledge, knowledge of anatomy, physiology, pathology, pathophysiology, microbiology, and pharmacology. In their view, biomedical knowledge puts constraints upon the ways in which patient characteristics, signs, and symptoms are related.

This hypothesized role of biomedical knowledge plays a central role in the discussion to follow. The position taken by Feltovich and Barrows in this debate can be paraphrased as follows: Comprehension and, therefore, the diagnosis of a case emanates from biomedical knowledge. Their point of view is supported by other investigators in the domain of medical diagnosis (e.g., Kuipers, 1984; Kuipers & Kassirer, 1984; Lesgold, 1984; Lesgold et al., 1988), who emphasized the role of biomedical knowledge in medical reasoning.

This perspective on diagnostic reasoning, however, was challenged by Patel, Evans, and Groen (1989), and others (e.g., Schmidt, Boshuizen, & Hobus, 1988), who suggested that medical *experts* predominantly use *clinical* knowledge instead of biomedical knowledge to represent and diagnose a patient problem.[2] According to these investigators, the application of biomedical knowledge is a particular characteristic of *nonexpert reasoning*.

Other authors, especially in the domain of artificial intelligence in medicine, take less extreme positions; for example, Szolovits, Patil, and Schwartz (1988) pointed out that biomedical knowledge may play a role in controlling the proliferation of hypotheses in clinical reasoning. One way to implement biomedical knowledge in a diagnostic AI program might be that knowledge is organized in multiple hierarchies, for example, anatomic and etiologic. An example is found in Pople's (1982) Caduceus. Another option is that biomedical knowledge is used to establish the effects of interactions among possible causes for patient findings. This approach was taken in Patil's (1983) ABEL.

### Experiments on the Role of Biomedical Knowledge in Clinical Reasoning

Not only are theories about the application of biomedical knowledge in medical diagnosis conflicting, the outcomes of empirical research in this area are paradoxical as well. Some investigators found an extensive use of biomedical knowledge in expert clinicians (e.g., Lesgold, 1984), whereas others reported virtual absence of references to biomedical knowledge in expert protocols (e.g., Schmidt et al., 1988). We will review these experiments in this section. Experimental designs and results are summarized in Table 1.

[2] Clinical knowledge is defined here as knowledge of attributes of sick people. It concerns itself with the ways in which a disease can manifest itself in patients, the kind of complaints one would expect given that disease, the nature and variability of the signs and symptoms, and the ways in which the disease can be managed. Biomedical knowledge, by contrast, concerns itself with the pathological principles, mechanisms, or processes underlying the manifestations of disease. It is phrased in terms of entities such as viruses or bacteria, in terms of tissue, organs, organ systems, or bodily functions (Patel et al., 1989).

TABLE 1
Overview of Experiments (Design and Outcomes)

| Researchers | Subjects | Cases | Method | Results |
|---|---|---|---|---|
| Patel, Groen, & Scott (1988) | *N*=24 8 beginning, 8 2nd-year, and 8 final-year students | Acute bacterial endocarditis | Post-hoc; biomedical science text read before diagnosing the case | Diagnostic performance increased with increasing levels of expertise |
| Patel & Groen (1986) | *N*=7 expert cardiologists | Acute bacterial endocarditis | Post-hoc | Pathophysiological explanations provided by experts were less extensive and more coherent than the students' explanations in Patel, Groen, & Scott (1988) |
| Patel, Evans, & Chawla (1986) | *N*=24 8 beginning, 8 2nd-year, and 8 final-year students | Acute bacterial endocarditis | Post-hoc; biomedical science text read after diagnosing the case | Diagnostic performance improved with increasing levels of expertise; expert pathophysiological explanations of the case were more coherent but contained fewer concepts |
| Patel, Arocha, & Groen, (1986) | *N*=4 2 cardiologists and 2 endocrinologists (clinicians and researchers) | Pericardial effusion; Hashimoto's disease | Post-hoc | Explanations provided by practitioners contained fewer biomedical concepts than researchers, especially in their own domain of expertise |
| Joseph & Patel (1990) | *N*=9 5 cardiologists and 4 endocrinologists | Hashimoto's disease | On-line | High-domain-knowledge subjects' problem representations were more coherent than low-domain-knowledge subjects', whereas fewer biomedical concepts were activated |

(*continued*)

TABLE 1 (Continued)

| Researchers | Subjects | Cases | Method | Results |
|---|---|---|---|---|
| Kaufman & Patel (1988) | $N=15$<br>5 final-year students, 5 residents, and 5 endocrinologists | Grave's thyrotoxicosis | Post-hoc[a] | Diagnostic accuracy increased with increasing levels of expertise; intermediates' explanations contained more biomedical concepts, whereas knowledge applied by novices contained more errors and inconsistencies |
| Schmidt, Boshuizen, & Hobus (1988) | $N=40$<br>8 beginning, 8 2nd-year, 8 4th-year, 8 6th-year students, and 8 internists | Acute bacterial endocarditis | Post-hoc | Inverted U-shaped relation between level of expertise and amount of biomedical knowledge applied |
| Lemieux & Bordage (1986) | $N=10$<br>4 poor and 5 outstanding students and 1 expert neurologist | Cervical arthrosis | On-line | Outstanding students applied more detailed biomedical knowledge than the expert; poor students' knowledge was either missing or represented as an unstructured list of facts |
| Lesgold et al. (1988) (2 studies; also partly reported in Lesgold, 1984) | (a) $N=23$<br>11 junior, 5 senior residents, and 5 expert radiologists | Lobectomy; Atelectasis; Multiple tumors (10 cases presented) | On-line (think aloud) | (a) & (b) Biomedical knowledge applied by experts seemed to be more precise and detailed and better suited to the presented cases, whereas novices' and intermediates' knowledge seemed to lead to many misperceptions |
| | (b) $N=12$<br>4 junior, 4 senior residents, and 4 expert radiologists | Same 3 cases as in (a) (5 cases presented) | On-line (think aloud & drawing | |

(continued)

TABLE 1 (Continued)

| Researchers | Subjects | Cases | Method | Results |
|---|---|---|---|---|
| Feltovich, Johnson, Moller, & Swanson (1984) (1 experiment, plus a deeper analysis) | $N$=12<br>4 students, 4 residents, and 4 expert pediatric cardiologists | Subvalvular aortic stenosis; Total anomalous pulmonary venous connection; Patent ductus arteriosus; Pulmonary atresia | On-line | Students' knowledge structures resembled in the textbook structure; in the experts' knowledge organization extra links, based on pathophysiological similarities, had been added; novices seemed to "reason through" biomedical knowledge in clinical reasoning |

[a] In addition, an on-line technique was applied in this study. Analysis of the on-line protocol focused on the dynamics of information acquisition and interpretation.

Many experiments in this domain have been done by the research group of Patel and Groen in Montreal, Canada (Joseph & Patel, 1987; Kaufman & Patel, 1988; Patel & Groen, 1986; Patel, Arocha, & Groen, 1986; Patel, Evans, & Chawla, 1986; Patel, Groen, & Scott, 1988). Their research goal has two aspects, namely, to investigate the nature of the knowledge applied, and the reasoning process itself. A typical experiment of this group consists of four elements: (1) presentation of a clinical case description; (2) recall of the case; (3) explanation of the signs and symptoms in terms of underlying processes; and (4) diagnosis (Patel & Groen, 1986). This experimental design is based on several assumptions. Patel and Groen assumed that an initial representation of a case was stored in the subject's working memory and was used to access relevant causal information. These two kinds of information were then combined to yield a diagnosis. Furthermore, they assumed that the initial representation could be probed by a free-recall task, whereas the knowledge applied could be accessed by asking the subjects to explain the process underlying the case (see Patel & Groen, 1986, pp. 93–94). In the major part of their experiments they applied this post-hoc approach considering pathophysiology protocols as traces of the knowledge activated.

The results of the experiments by Patel and Groen and their colleagues suggest a rather complex relationship between the knowledge their subjects applied and the outcomes of the diagnostic process. Novices appeared to construct incoherent representations of problems presented to them and made factual mistakes when describing the processes underlying a case. Furthermore, the amount of clinical experience appeared to be negatively related with the number of biomedical concepts used in the explanations. For instance, students, clinical researchers, or clinicians diagnosing cases in an unfamiliar domain applied more biomedical knowledge than clinical experts in their own domain (Kaufman & Patel, 1988; Patel, Arocha, & Groen, 1986; Patel, Evans, & Chawla, 1986; Patel & Groen, 1986; Patel et al., 1988).

However, some of these studies do not allow for general conclusions about differences in the amount of biomedical knowledge applied by subjects with different levels of expertise. In two of the three studies involving students, extra information concerning relevant biomedical concepts was provided (Patel, Evans, & Chawla, 1986; Patel et al., 1988). In the experiment by Patel et al. (1988), students first read three texts about relevant biomedical subjects (microcirculation, physiology of fever, and human hemodynamics) before diagnosing the criterion case of acute bacterial endocarditis. This part of the experimental procedure may have affected the level of detail and the nature of the concepts applied in the pathophysiological explanations provided by the students, which might account for the observed differences between experts and students.

Nevertheless, in a study by Schmidt et al. (1988), which used a more rigorous methodology, the same phenomenon was found. These authors replicated the Patel and Groen (1986) study using the same clinical case and 40 subjects at five levels of expertise, but did not provide the additional biomedical texts. They found an inverted U-shaped relationship between level of expertise and number of biomedical propositions applied: 4th- and 6th-year medical students produced more propositions of a biomedical nature than laypersons and 2nd-year students. The experts, however, produced the fewest biomedical propositions.

Most investigations discussed so far applied a post-hoc method based on a reactivation of the knowledge applied after a diagnosis was produced. The remainder of this section will be dedicated to studies using an on-line approach to diagnostic reasoning that attempt to "tap" the knowledge applied while the diagnostic process is in progress.

Lemieux and Bordage (1986) were interested in differences in knowledge structure that might account for performance differences among students. Their results showed that the outstanding students applied elaborate biomedical knowledge, whereas the poor students either applied unrelated lists of facts, or missed the relevant knowledge. The think-aloud protocol of an expert acquired in order to provide a frame of reference was far less detailed in terms of biomedical concepts than the protocols of the outstanding students. Lesgold and colleagues (Lesgold, 1984; Lesgold et al., 1988) investigated the development of expertise in diagnosing X-rays. Lesgold et al. explored the relation between diagnostic outcomes and the characteristics of the knowledge applied. They found large differences among subjects in terms of structures attended to and concepts applied while describing X-rays. Their findings suggest that with increasing levels of expertise, the biomedical knowledge applied becomes more detailed and more tailored to the case at hand.

Feltovich, Johnson, Moller, and Swanson (1984) also investigated the relation between level of expertise and the structure of the knowledge applied. Their results indicate that increasing levels of expertise are associated with changes in knowledge structures pertaining to the mutual relations between diseases. Student knowledge structures are apparently organized in a hierarchical fashion, reflecting the textbook organization (e.g., cyanotic and noncyanotic congenital heart diseases), but this strict hierarchy is disrupted in the expert group, where new links between disorders with the same underlying pathophysiology and symptomatology appear. Hassebrock and Prietula (1986), two other investigators from this group, reported more detailed descriptions of the knowledge applied in diagnosis by subjects with different levels of expertise. It appeared that students applied elaborate biomedical knowledge "reasoning through" a biomedical conceptual network in order to activate a relevant diagnosis. Experts tended to recognize specific findings as indicators of a general class of heart defects.

The experiments reviewed are summarized in Table 1. This overview shows that the methodology applied, the nature of the cases presented, the levels of expertise studied, and the conclusions are highly diverse. The most striking differences are found between the studies conducted by Patel and colleagues and those of others. Patel and colleagues used a post-hoc methodology in five of six investigations, whereas in the other studies, on-line methods prevailed. In addition, the study by Schmidt et al. (1988), which largely replicated and extended Patel's findings, also used a post-hoc approach. One is tempted to conclude that the contradictions among findings may be due to differences in methodology. However, Patel's findings are at least partially supported by the on-line studies by Lemieux and Bordage (1986) and Hassebrock and Prietula (1986).

On the other hand, some of the studies suggesting a major role for biomedical knowledge in medical expert diagnosis, provide only circumstantial evidence for this proposition. This is particularly the case in Feltovich et al. (1984). The studies by Lesgold and colleagues, however, provide unequivocal support for the thesis that the development of expertise in medicine is characterized by an increasingly flexible and more adequate use of biomedical concepts. Their experts clearly showed indications of a deeper understanding of the nature of diseases, and this understanding was based on more elaborate schemata consisting of anatomical and pathophysiological concepts. This may, however, be an idiosyncratic property of the radiology domain (Boshuizen & Schmidt, 1991).

In conclusion, the experiments reviewed do not provide a sufficient basis to decide whether biomedical knowledge plays a major role in expert medical diagnosis, and if so, *how* this role is played. Neither can major conclusions be drawn about its role in preceding stages of development. Our first and tentative conclusion, therefore, is that the overt application of biomedical knowledge decreases with increasing levels of expertise. In addition, it is suggested that in the course of the development, biomedical and clinical knowledge are subject to changes, both in structure and function. It is, however, unclear which mechanisms induce those changes.

In order to investigate the role of biomedical knowledge in medical diagnosis, two experiments will be presented. The aim of the first experiment was to investigate whether the findings by Patel and colleagues (Kaufman & Patel, 1988; Patel, Evans, & Chawla, 1986; Patel & Groen, 1986; Patel et al., 1988) using a post-hoc methodology, could be replicated utilizing an *on-line method*. Four subjects with different levels of expertise were extensively studied while reasoning about a case presented to them in a sequential fashion. A second purpose of this on-line study was to elucidate the role of biomedical knowledge in clinical reasoning and to suggest hypotheses about the mechanisms that may be responsible for changes in the course of the development toward expertise. The second experiment was designed to test these hypotheses.

## EXPERIMENT 1

### Method

Four subjects at different levels of medical expertise participated in the study: One 2nd-year medical student; a 4th-year student who had nearly finished preclinical training; a 5th-year student who had finished both a primary care and an internal medicine internship, and a family physician with 4 years of experience. The 2nd-year student was the novice in this study. The 4th- and 5th-year students were Intermediate 1 and Intermediate 2. The family physician was the expert.

The subjects were presented with a case of a 38-year-old, unemployed man with a history of neurotic depressions and alcohol abuse. One year earlier, this patient had had an attack of pancreatitis, and now called the family physician with a complaint of severe, boring pain in the upper part of the abdomen. He suffers from a chronic relapsing alcohol-induced pancreatitis with minor pancreatic insufficiency. The symptoms associated with this disease and the underlying pathophysiological mechanism are described in Appendix A. The case was presented on 48 typed cards, each containing one or more items of information that characterized the patient: history taking, physical examination, and lab findings (the case items are represented in Appendix B).

The subjects were asked to think aloud while being presented with the cards in a sequential fashion and to provide a differential diagnosis at the end. These sessions were tape recorded and verbatim transcripts were produced.

### Analysis

*Think-Aloud Protocols.* The analysis of the think-aloud protocols aimed to identify those parts of the protocols in which biomedical and clinical knowledge were applied to diagnose the case. The identification of these parts was achieved in a step-by-step procedure.

The first step was segmentation, based on pauses in the protocols. Next, general comments and meta-statements were removed, such as "I would leave this point aside and first concentrate on...," "I would never have asked that question!" or "I am not sure whether what I am saying now is really right."

In the next step, the remaining segments were rewritten as propositions consisting of a relation[3] and a set of arguments (Anderson, 1985). A major part of the arguments in the propositions consists of a case item, an interpretation of a case item, or a hypothesis generated. These propositions were represented in a semigraphical way: [argument→(relation label) argument]. For example, the proposition [appetite: poor→(cond) liver] consists of a conditional relation and the arguments "appetite: poor" and "liver"; the proposition [BP: 145/90→borderline] consists of an unlabeled relation and the arguments "BP: 145/90" and "borderline."

As a final step the propositions were classified as biomedical or clinical. This distinction was based on the object of the proposition. Propositions concerning pathological principles, mechanisms, or processes underlying the manifestations of disease were classified as biomedical propositions. They are phrased in terms of entities such as viruses, bacteria, stones, or carcinomas, or in terms of tissue, organs, organ systems, or body functions. A proposition in which the name of an organ was used, was only classified as "biomedical" if it contained a location link (e.g., "is adjacent to . . .") or a specification of the biomedical process operating on that organ (e.g., "chemical irritation of . . .")[4] [Peritoneal irritation→(cau) automatically decreases bowel motility] is an example of such a proposition. By contrast, propositions concerning attributes of people, including their diseases, are labeled clinical (Patel et al., 1989). These propositions are concerned with the ways in which a disease can manifest itself in the patient, the kind of complaints one would expect given a specific hypothesis, the nature and variability of the signs and symptoms, and the ways in which the disease can be managed. As the classification principle is based on the object of a proposition, often propositions from adjacent protocol fragments must be taken into account. Thus, an interrater agreement of .95 was obtained. When raters disagreed, items were discussed. If no agreement could be reached, the items were classified as "clinical." An example of the analysis of a protocol

[3] Relations can be *causative* (e.g., peritoneal irritation *automatically decreases* the bowel motility); *conditional* (e.g., boring pain in the upper abdomen *is an indication for* pancreatitis); *specification* (e.g., something is wrong with the peritoneum, *probably* a beginning peritonitis); *is-a* (=*abstraction*) (e.g., he has had this sort of pain more often recently, that *is a* chronic process manifesting itself from time to time); *identity* (e.g., pancreatitis *is* an inflammation of the pancreas), *negation* (e.g., most of the time it vanished within a day's time, *evidence against* a stomach rupture); and *location* (e.g., I must also think of the gallbladder, which *is close to* the pancreas). Very often the relation is not specified (e.g., in the responses "pancreatitis" or "that is quite high" to the lab finding of serum amylase 128E). In those cases, the relation is represented as an unlabeled link.

[4] Nonspecific or lay terms (e.g., "damage of" or "assault on") used to describe the process were not regarded as referring to a biomedical process.

part is given in Appendix C. The biomedical propositions were counted and this number was divided by the total number of propositions extracted.

**Results**

*Characteristics of the Think-Aloud Protocols.* The four think-aloud protocols largely varied in elaborateness. The longest protocol was produced by the expert and consisted of 256 sgments, from which 78 knowledge-application propositions could be extracted. Intermediate 1's protocol was the shortest, consisting of 120 segments, containing 71 propositions. Intermediate 2's protocol consisted of 135 segments and, though not the shortest of the four, only 39 propositions could be extracted. The novice's protocol contained 160 segments, including 75 propositions. As we have only 1 subject per level of expertise, it is impossible to decide whether these differences are characteristic for the experimental subjects or whether they are related to group differences.

*Application of Biomedical Knowledge.* The number of biomedical propositions that could be identified in these sets of knowledge-application propositions was not proportional to the total number of propositions: The protocols of the 2 subjects at a lower level of expertise contained more biomedical propositions, proportionally and in an absolute sense, than the other two protocols.

Because the total number of propositions produced varied among subjects, proportions of biomedical propositions per subject were computed. Figure 1 shows that the proportions of biomedical propositions decreased with increasing levels of expertise. More than 50% of the 2nd-year student's propositions were labeled as biomedical, monotonically decreasing to less than 10% in the expert's protocol. This result, by and large, replicates the findings reported by Patel and colleagues using a post-hoc procedure, and suggests that the overt role of biomedical knowledge in the development of a mental representation of a medical case decreases with an increasing level of expertise.

A qualitative analysis[5] of the propositions extracted from the think-aloud protocols shows that the quality of the applied biomedical knowledge also changes over time. In the novice's protocol many propositions contained invalid concepts and/or relations. An example of the application of an invalid concept is found in the response to Item 23, that according to other people, the patient has been jaundiced.

> ..Well that suggests a liver disease..err..in which case..pirubin I believe can cause yellowing of the skin.

[5] We wish to thank P.P.M. Hobus for this analysis (personal communication, 1989).

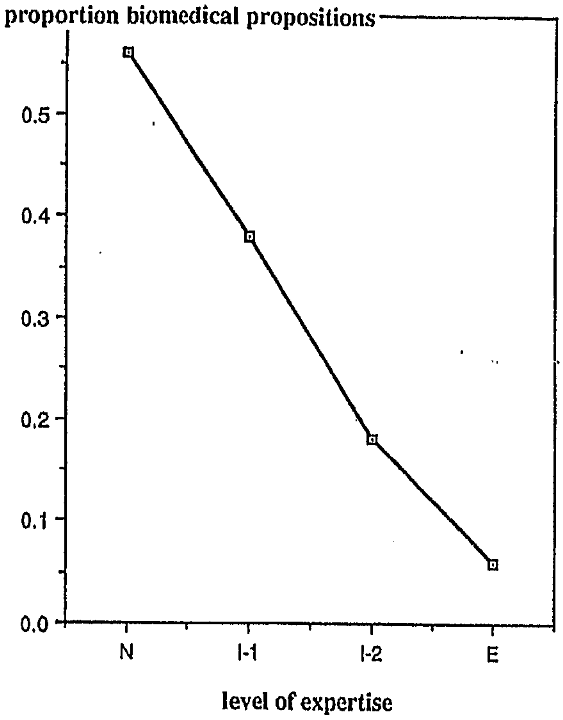


**Figure 1.** Proportion of propositions that were classified as biomedical propositions extracted from the think-aloud protocols of 4 subjects with different levels of expertise.

The novice applied the term "pirubin" in order to explain the case finding of "occasional jaundice." Probably it is a corruption of "bilirubin," which is produced in the liver. An example of the application of an invalid relation can be found in this subject's response to Item 48a, the high level of serum amylase:

> Amylase, that's starch I believe.

The novice incorrectly equated the enzyme amylase with starch, the nutritional element that is digested with the help of amylase.

No such obvious mistakes and misconceptions occurred in the other subjects' protocols. Hence, the relative share of biomedical propositions in the total number of propositions diminished with an increasing level of expertise, whereas the accuracy of the biomedical knowledge applied was better in the higher levels of expertise.

Another interesting phenomenon concerns Intermediate 1's style of reasoning, a style not encountered in the other subjects' protocols. This style can be characterized as a detailed, step-by-step approach from the case signs and symptoms to the final diagnosis. A typical example of such a line of reasoning was found in the response to Item 34, the patient's pulse rate.

> ...The past two symptoms.together.mean that there's no inflammation.. and that would eliminate a..a..cholecystitis..and would rather mean an.. obstruction of the biliary tract.caused by a stone, for instance.or, what may be the case too, by a carcinoma, but I wouldn't..although, it might be possible, lost 5 kilograms in weight.

Intermediate 1 first concluded from the patient's normal temperature and pulse rate that there is no inflammation and hence no cholecystitis. Then he wondered what other kind of process could cause biliary tract obstruction, his main hypothesis. As a first solution, a stone was thought of, but a carcinoma was considered as well. Confirming evidence for this last possibility was found in the patient's weight loss. The number of biomedical concepts is high in this line of reasoning (inflammation, obstruction of the biliary tract, stone, and carcinoma). Apparently, this subject generated potential explanations for a set of symptoms, based on biomedical knowledge. Confirming evidence is then found in another symptom presented earlier.

Noticeably, this approach differs widely from the expert's approach, which is far less detailed. More than any other subject, the expert took the patient's background into account in generating and verifying or falsifying hypotheses. For example, in response to Item 29 concerning the patient's weight loss in the past 6 months the expert stated:

> ...This might mean two things: He's an excessive drinker..eats irregularly.. has a poor appetite..what food he takes seems a bit deficient, judging from what he describes, so it may be quite possible that his food intake in insufficient on the one hand; ..he has had an acute pancreatitis in the past, a drinking habit, fatty liver, erm. He's only er...he's only in his thirties..of course it might be a manifestation of a process in the liver..or a malign process in the pancreas; two tracks remain open from this information.

Based on information about the patient's history, the expert concluded that the patient's weight loss may be caused either by insufficient food intake or by a malignancy, although the patient is only in his 30's (and hence, rather young for such a disease). Traces of this kind of reasoning were also found in Intermediate 2's protocol, where indications of biomedical reasoning were almost absent.

Taken together, these results indicate a marked shift from the application of biomedical reasoning toward clinical knowledge. In this study, the transition from the application of biomedical knowledge to the application of clinical knowledge seems to be associated with the transition from preclinical to clinical education, because Intermediate 1 was about to enter internships, whereas Intermediate 2 had already had 9 months of experience in the clinic and in general practice. Hence, practical experience might play an important role in this change.

### Discussion

The first objective of Experiment 1 was to replicate the findings by Patel and colleagues (Joseph & Patel, 1990; Kaufman & Patel, 1988; Patel, Arocha, & Groen, 1986; Patel, Evans, & Chawla, 1986; Patel & Groen, 1986; Patel et al., 1988) and Schmidt et al. (1988) with an on-line method. The results presented so far agree with the findings by these investigators, who showed that increasing levels of expertise are associated with a decreasing application of biomedical knowledge using a post-hoc method. They are at variance with the findings of those researchers who claim an important role for biomedical knowledge in expert clinical reasoning (Feltovich et al., 1984; Lesgold, 1984; Lesgold et al., 1988). Our results suggest that the conflicting research outcomes found in the literature cannot be attributed to differences in the method applied.

The second objective of this first study was to clarify the role of biomedical knowledge in clinical reasoning and to suggest hypotheses about the mechanisms that may be responsible for changes in the course of development toward expertise. The results here suggest a rather abrupt shift from the application of biomedical knowledge to clinical knowledge. A peak in the functionality of the biomedical knowledge applied was observed at the lower intermediate level. In the protocols of the two subjects with practical experience, the application of clinical knowledge was prominent. Those results suggest that biomedical knowledge is acquired and used as the major instrument in interpreting clinical information in the early stages of development. After misconceptions have been removed, biomedical knowledge provides a reliable tool for forming a coherent mental representation of a clinical case as the protocol of Intermediate 1 suggests. However, this application of biomedical knowledge does not seem to endure. The transition to the next stage seems to be initiated by *the effect of practical experience*. In this next stage, the application of biomedical knowledge appears virtually absent. Instead, clinical knowledge appears to be predominantly applied.

Yet, the question remains: Which mechanism underlies this shift in knowledge application? Some candidates might be considered. The first hypothesis is very straightforward: An increase in the application of knowledge reflects an increase in the availability, whereas a decrease in application reflects a decrease in availability. This would imply that biomedical knowledge is acquired and applied at the early stages of the intellectual development of physicians-to-be. This kind of knowledge may initially be advantageous in the students' understanding of clinical phenomena. However, after experiences with real patients and their diseases, biomedical knowledge might become gradually less important and may eventually become inaccessible or rudimentary; even retroactive interference might play a role in this process. Clinical knowledge, on the other hand, continuously increases. The work of Ackermann and Barbichon (1963) shows that the acquisition of expertise in other domains (chemical and electrical engineering) is associated with such a forgetting of details.

A second hypothesis explaining the decrease in application of biomedical knowledge is that biomedical knowledge is still available to medical experts, but does not play a role in their clinical reasoning. It is inert. Clinicians may tend to use experiential knowledge instead. This experiential knowledge is acquired as a result of extended practice and continued exposure to the many different ways in which disease manifests itself. The inert biomedical knowledge would still be accessible, but is of such a nature that it is not activated in medical diagnosis. Leinhardt (1987) demonstrated such a phenomenon in the domain of teaching. In teacher education, mathematics is considered very important for the teaching of elementary arithmetic. Leinhardt found that this knowledge can be remembered over more than 20 years, but it is not integrated in the teachers' experiential, situated knowledge of teaching elementary arithmetic. Keeping the analogy with Leinhardt's results, this hypothesis implies that biomedical knowledge will not be helpful in the tasks that medical students have to perform after entering internship; instead, clinical knowledge is acquired and applied. As a consequence, knowledge applied in clinical reasoning and biomedical knowledge would "represent two different worlds," with no relation at all (Patel et al., 1989).

The third hypothesis explains the decrease in the (overt) application of biomedical knowledge as a result of its increasingly tacit and automatic nature. It assumes that biomedical knowledge stays available to the medical expert, but it no longer necessarily plays a prominent role in clinical reasoning. Schmidt and Boshuizen (in press) suggested that this tacit application of biomedical knowledge by medical experts might have resulted from a process they called "knowledge encapsulation." They observed that case recall and post-hoc explanation by advanced students improved dramatically when they were allowed to activate relevant biomedical knowledge before or during case processing. Laypersons and experts did not benefit from ample activation time, nor were they hampered by extreme time constraints. They concluded that the intermediates actively retrieved knowledge and had to reason through these knowledge structures, whereas expert knowledge was more readily available, and its activation effortless.

Schmidt and Boshuizen (in press) assumed that the repeated application of biomedical knowledge in clinical reasoning at the earlier stages of development toward medical expertise resulted in the subsumption of lower level, detailed propositions under higher level, sometimes clinical propositions. This encapsulation process is thought to have resulted in easily accessible and flexible knowledge structures with short search paths. This hypothesis is in agreement with results by Patel, Arocha, and Groen (1986) who found that experts facing familiar problems applied less, and less detailed, biomedical knowledge than experts facing unfamiliar problems. Their results also suggest that biomedical knowledge stays available and can be retrieved whenever necessary. This necessity might occur in more complex cases or when an explanation is asked for (cf. ABEL's knowledge multilevel link structure; Patil, 1986). In those cases it is assumed that encapsulated knowledge can be easily expanded. For example, the direct link between drug abuse and possible endocarditis made by one expert subject (No. 5) in Patel and Groen (1986) probably can be easily expanded to a chain of propositions [drug abuse → (cond) contaminated needles → (cau) introduction of bacteria into the circulation → (cau) sepsis → (cau) infection of the endocardium → (iden) endocarditis].[6] In unambiguous cases, such as the one used in Experiment 1, application of biomedical knowledge would occur fairly implicitly, therefore, leaving hardly any verbal traces in the think-aloud protocols (Ericsson & Simon, 1984).

## EXPERIMENT 2

In Experiment 2 these following three hypotheses about the role of biomedical knowledge in medical diagnosis at successive levels of expertise were investigated:

1. After a certain stage in the development toward expertise, biomedical knowledge becomes rudimentary.
2. After a certain stage in the development toward expertise, biomedical knowledge becomes inert.
3. During the development toward expertise, biomedical knowledge becomes encapsulated in clinical knowledge.

Hypotheses 1 through 3 equally predict that the think-aloud protocols of subjects of increasing levels of expertise will reveal a decrease in the application of biomedical knowledge. However, Hypothesis 1 also predicts a decrease in the biomedical knowledge applied post-hoc. If expert biomedical knowledge has become rudimentary, the experts will be less able to explain the pathophysiology underlying the case. Hypothesis 2 predicts no such

[6] The reader may observe that the notion of knowledge encapsulation is analogous to the notions of knowledge compilation (Anderson, 1983; 1985; 1987), knowledge restructuring and tuning (Rumelhart & Norman, 1978), and chunking (Laird, Rosenbloom, & Newell, 1986), all leading to a marked acceleration of search through a knowledge base. The new term was, however, introduced to emphasize the abbreviations in search processes resulting from knowledge application without rendering the original, detailed knowledge base inaccessible.

decrease, nor does Hypothesis 3. Thus, in order to discriminate between Hypotheses 2 and 3, an extra measure is needed. This measure can be found in the correspondence between the knowledge applied in clinical reasoning and the explanation of the case provided post-hoc. If medical experts apply encapsulated knowledge in clinical reasoning, then Hypothesis 3 predicts that they will expand this encapsulated knowledge when asked to explain the case. Hence, propositions in the think-aloud protocols of medical experts will be reflected in longer chains of propositions with coinciding beginning and ending arguments in the post-hoc explanations of the case. Novices' and intermediates' knowledge is not yet encapsulated. Hence, their think-aloud protocols will include fewer abbreviations. On the other hand, if expert biomedical knowledge is inert, as Hypothesis 2 assumed, and is not applied in clinical reasoning, and if biomedical and clinical knowledge are two worlds apart, then no such increase in number of abbreviations will be expected.

**Method**

*Subjects.* In Experiment 2, 20 subjects participated. Six subjects were 2nd-year students, having the same level of expertise as the novice in Experiment 1 (novices). Four subjects were at the same level of expertise as Intermediate 1, namely, at the end of the 4th-year (I-1's). Five 5th-year subjects (Intermediate 2 level), who had finished their internships in internal and family medicine (I-2's), participated. The expert group consisted of 5 family physicians with about 4 years of experience (experts).

*Material and Procedure.* The subjects were presented with the same pancreatitis case as was used in Experiment 1. The subjects' task was to diagnose the case while thinking aloud. After completing the case they were asked to describe (in writing) the pathophysiological processes that underlie the case. Subjects were tested individually.

*Analysis.* The think-aloud protocols were analyzed with respect to the application of clinical and biomedical knowledge. The same procedure was applied as in Experiment 1. One audio recording (of Subject 5-12, a 5th-year student) contained so much noise that no transcription could be derived from it. Therefore, analyses of the think-aloud protocols were based on the data of 19 subjects.

The explanations of the underlying pathophysiological process were analyzed using a method described by Patel and Groen (1986), who segmented these texts into propositions consisting of two arguments and a relation. The propositions were counted and represented as a semantic network.

Finally, the correspondence was assessed between the propositions derived from the think-aloud protocols and the propositions derived from the patho-

TABLE 2
Summary Table Descriptors of the Think-Aloud Protocols of Subjects with Four Different Levels of Expertise (Means and Standard Deviations)

| | Novices[a] M (SD) | I-1's[b] M (SD) | I-2's[b] M (SD) | Experts[c] M (SD) |
|---|---|---|---|---|
| No. of propositions | 38.8 (15.4) | 46.0 (4.69) | 25.0 (8.04) | 35.8 (11.56) |
| No. of biomedical propositions | 11.0 (8.60) | 21.8 (12.55) | 5.0 (4.83) | 5.2 (2.49) |
| Proportion of biomedical propositions | .24 (.18) | .47 (.26) | .17 (.13) | .14 (.06) |

[a] $n=6$. [b] $n=4$. [c] $n=5$.

physiological explanations. Arguments served as anchor points in this analysis. First, arguments of the think-aloud propositions were matched to the post-hoc propositions. Next, the relations between the matching arguments were classified. Three kinds of relations can be distinguished:

1. Identical relations (the relations between the arguments matched are identical).
2. Abbreviations (the proposition in the think-aloud protocol is an abbreviation of a chain of at least two propositions in the post-hoc explanation).
3. Other relations.

Increasing knowledge encapsulation will show itself in an increasing number of abbreviations.

The data were analyzed using SPSS-X MANOVA. Polynomial contrast analyses were completed in order to test for nonlinearity of the relations between dependent variables and the subjects' levels of expertise.[7]

## Results

*Think-Aloud Protocols.* Table 2 shows that the number of biomedical propositions extracted from the think-aloud protocols is associated with level of expertise, $F(3, 15)=4.102$, $p<.05$. As individual differences in length of the think-aloud protocols were substantial (especially in the 2nd-year students and in the expert group), the number of applied biomedical propositions were also expressed as a proportion of the total number of propositions extracted from the protocols, $F(3, 15)=3.199$, $p<.06$.

[7] This analysis calculates linear, quadratic, and cubic terms in the between-groups variance. Nonlinear relations result in significant deviations from linearity and a significant quadratic and/or cubic term without significant deviations.

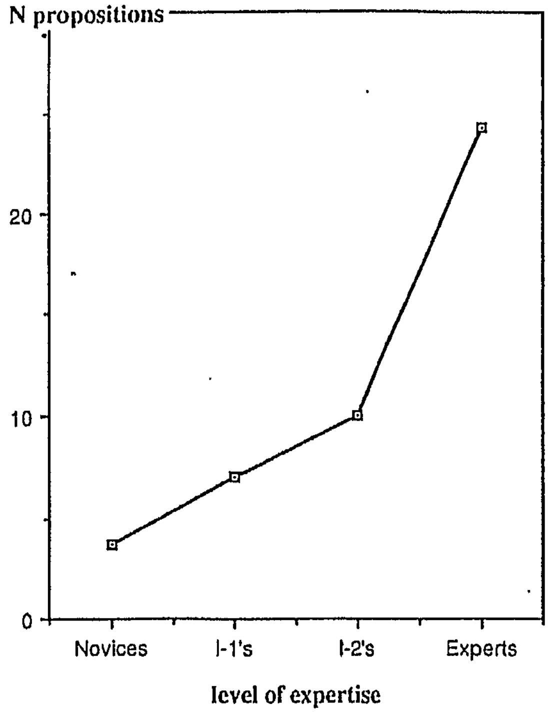


**Figure 2.** Number of propositions applied by subjects of different levels with expertise in explaining the pathophysiological process that caused a patient's complaints.

Apparently, the top of the curve is to be found in the 4th-year students' group. About half of the knowledge they applied in the think-aloud protocols was derived from the biomedical sciences. Associated with this curve is a significant cubic component, $F(1, 15) = 6.215$, $p < .05$. The linear and quadratic components of the curve have $p$ values greater than .10. These results confirm the finding that, contrary to intermediates, experts seem to apply hardly any biomedical knowledge. However, the monotonic decrease that was found in Experiment 1 could not be replicated.

*Post-hoc Pathophysiological Explanations.* The number of propositions used to explain the pathophysiological process underlying the patient's complaints is associated with expertise level, $F(3, 16) = 3.567, p < .05$. Polynomial analysis shows that only a linear component in this relation is significant, $F(1, 16) = 10.443, p < .05$, without significant deviations from linearity ($p > .25$); see Figure 2.. This result suggests that the experts' biomedical knowledge has not become rudimentary. On the contrary, increasing levels of expertise are associated with a monotonic increase in the knowledge of the biomedical sciences. These data clearly contradict the hypothesis that

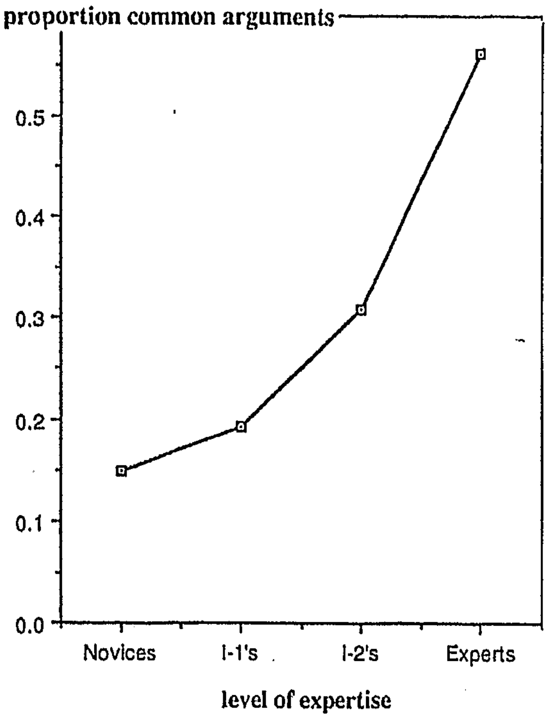


**Figure 3.** Proportion of common arguments in the propositions derived from the think-aloud protocols and the pathophysiological explanations.

the use of biomedical propositions decreases with expertise due to rudimentariness or inaccessibility of such knowledge.

*Correspondence Between Think-Aloud Protocols and Post-hoc Explanations.* The two remaining hypotheses predict differences in correspondence between think-aloud protocols and post-hoc explanations. The results from the correspondence analyses are represented in Figures 3 and 4. Figure 3 shows that the proportion of arguments in the post-hoc explanation propositions corresponding to an argument in propositions extracted from the think-aloud protocols increased monotonically with increasing levels of expertise, $F(3, 15) = 19.887$, $p < .0001$. Among the novices, argument correspondence was only 14.9%. This percentage increased to 56.2% in the expert group. The linear component associated with this trend is highly significant, $F(1, 15) = 57.903$, $p < .0001$, without significant deviations from linearity ($p > .25$).

Proposition-mapping analysis showed the following results: Analysis of the number of propositions in the think-aloud protocols identical with the pathophysiological networks shows no differential effects of expertise level,

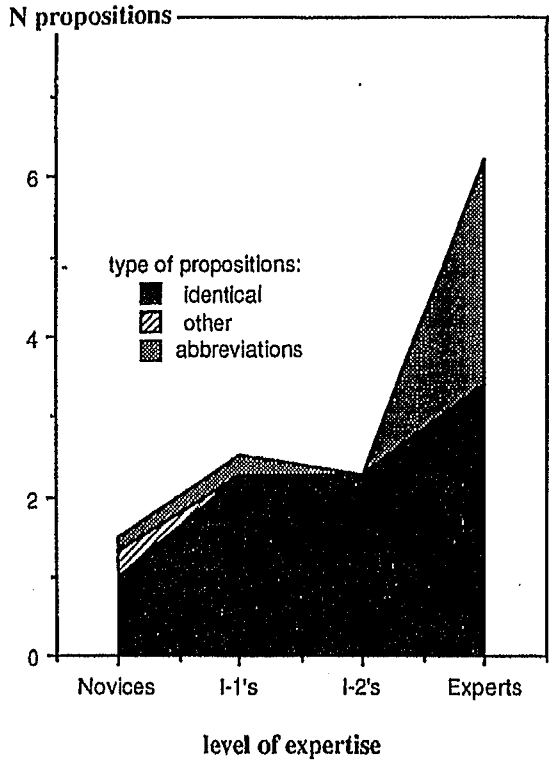


**Figure 4.** Three types of propositions representing the correspondence between on-line and post-hoc propositions generated by subjects with four different levels of expertise.

$F(3, 15) = 1.63$, $p > .20$. The same was found for the number of "other" propositions, $F(3, 15) = 1.71$, $p > .20$. By contrast, the number of abbreviations turned out to be significantly related to level of expertise, $F(3, 15) = 9.01$, $p < .001$. There is a significant linear component in this effect, $F(1, 15) = 21.81$, $p < .001$, without significant deviations ($p > .05$). Figure 4 represents the data.

These results support Hypothesis 3 because only the number of abbreviations increased with level of expertise. Expert biomedical knowledge seems to be encapsulated under clinical propositions.

**Discussion**

Experiment 2 served two purposes: (1) to replicate Experiment 1 with more subjects, and (2) to investigate three hypotheses that were proposed in order to explain the phenomena observed in Experiment 1. The outcomes of the analysis of the think-aloud protocols largely replicate the results of Experiment 1, with one exception. In Experiment 1 the novice applied (absolutely and proportionally) most biomedical propositions, although the contents of these propositions often revealed misconceptions. In Experiment 2, the subjects at the lower intermediate level applied the most biomedical propositions, reflecting the amount of knowledge of students at these stages. Experiment 2 suggests that the application of biomedical knowledge first increases and later again decreases with increasing levels of expertise. Analysis of the post-hoc explanations and the correspondence between think-aloud protocols and post-hoc explanations led us to the conclusion that this later decline in the application of biomedical knowledge did not result from decay of biomedical knowledge leading to a rudimentary state in expertise, nor was it a result of inertia of biomedical knowledge in diagnostic reasoning. Instead, biomedical knowledge plays its role in a tacit way as it is encapsulated in clinical knowledge.

Qualitative aspects of the subjects' lines of reasoning and explanations further validated this conclusion. Comparison of the subjects' post-hoc explanations with the canonical explanation as provided in Appendix A revealed that no novice described this canonical process. Only one novice tried to explain the underlying process in some detail, but this subject assumed that a kind of poisoning process caused the liver (or another organ) to become inflamed. The I-1s' explanations were far more detailed, but again, no match for the canonical explanation could be found. Two of the I-2 subjects' explanations contained arguments that were similar to the canonical explanation. Finally, the experts' explanations show again more detail, and the concepts applied were more similar to the concepts in the canonical explanation. For example, in the experts' semantic networks, concepts such as "destruction of tissue," "release of pancreatic enzymes," "diminished digestion," "atrophy of the liver cells," "atypical inflammation reaction," "scarring of tissue," "changed structure of the pancreatic duct," "sensibilization for other toxic and physical influences," and "deposit of metabolites" were found.

Hence, it can be concluded that not only the elaborateness of the pathophysiological explanations provided increased, as was illustrated in Figure 3, but the explanations became increasingly similar to the canonical explanation described in Appendix A. Both findings indicate that the biomedical knowledge used by our subjects in explaining the case expands with increasing levels of expertise.

## GENERAL DISCUSSION

After having answered the questions of whether the overt application of biomedical knowledge decreases with increasing levels of expertise and what kind of process underlies this change, two questions remain: (1) What kind of learning process results in knowledge encapsulation; and (2) What is the generality of the results? These questions will be addressed in the following.

**Learning Processes**

Regarding the question about learning processes resulting in knowledge encapsulation, two possibilities can be discerned: (1) encapsulation is a direct result of problem solving, namely, diagnosis; or (2) encapsulation is an indirect result. In order to decide between the two candidates it is necessary to analyze the situation in which this process takes place. When students first enter clinical internships they have a great deal of theoretical knowledge: They know about all kinds of biomedical mechanisms, normal functions, how disturbances may result in diseases, and how the body reacts to counter these disturbances to bring the situation back to normal. Furthermore, they have knowledge about specific disturbances and specific diseases and how medical treatment can be used to fight these specific ailments (and about the mechanisms underlying their effectiveness). During the preclinical period, students see few if any patients. Then, upon entering internship, things suddenly change. Now, a great part of their time is spent with patients. They are taken on ward rounds and hear patients described and explained by their supervisors. They are expected to take patients' histories, do physical examinations, and make diagnoses all by themselves. Such internships serve different educational objectives. Among others, students must learn to apply the theoretical subject matter studied in the preclinical period, and it gives them the opportunity to see, hear, smell, and feel for themselves the concepts they only know from books and audiovisual media.

What effect can this new situation, these new tasks, and these new experiences have on the students' knowledge structures? The answer to that question will largely depend on the way students approach these new tasks. One possibility is to stick to the reasoning and knowledge application procedures they used to apply in the preclinical period: Active application of biomedical and theoretical knowledge as demonstrated by the 4th-year students in both experiments. Analogous to the knowledge-compilation processes described by Anderson (1983, 1987), repeated knowledge application of this kind might result in abbreviations of search paths (see also Elio & Scharf, 1990) leading to knowledge encapsulation. Verbal reports of students about the time pressure experienced during internship suggests, however, that this possibility is not very plausible. They simply do not have time to reason through an elaborate causal network in order to find a diagnosis: Physical examination and history taking must always be efficient, responses must always be quick.

An alternative possibility is that during the internships, students acquire new clinical and experiential knowledge directly linking patient signs and symptoms to diagnostic hypotheses, and apply this kind of knowledge in their clinical reasoning. Such a process would result in a shift from biomedical to clinical knowledge application, and, if nothing else happens, it also would lead to inertia of the biomedical/theoretical knowledge base. Hence, the encapsulation of biomedical knowledge under higher order propositions observed in this study must emanate from an additional process invoking the integration of biomedical and clinical knowledge, of theoretical and experiential knowledge.

The latter explanation of the phenomena observed bears close resemblance to what Collins and colleagues (Collins, 1990; Collins, Brown, & Newman, 1989) described as the acquisition of a robust knowledge base. Such a robust knowledge base results from the active integration of general and situated knowledge: (1) articulating a global framework (the biomedical knowledge); (2) reflecting on situated experiences (the clinical experiences) as they relate to the global framework; and (3) exploring in order to elaborate connections between situated knowledge and the global framework. Collins (1990) used the example of learning to find your way in a strange city. The best way, he states, is neither map study nor plain experience, but the active matching of experiential knowledge to relevant general knowledge. In the acquisition of robust medical knowledge, this active matching of biomedical and experiential knowledge might result from self-explanation, from restudying biomedical sciences, and from discussions with their fellow students and during staff meetings.

A compilation-like process would result in a gradual decrease of the application of biomedical knowledge in clinical reasoning associated with a gradual increase in the number of abbreviations. No such phenomenon was observed. Instead, a rather abrupt shift from biomedical to clinical knowledge application has been found, whereas the associated appearance of abbreviations was delayed until later stages of development. This suggests that knowledge encapsulation is an indirect effect of practical experience, needing an extra cognitive effort by the apprentice physician.

### Generality

The applicability and generality of the research outcomes and conclusions here include several aspects: the kind of problem used; the specific characteristics of the curriculum of the medical school where our subjects were recruited; and finally, the kind of expertise medical expertise is an example of.

The problem used, alcohol-induced pancreatitis, is a rather rare disease with an incidence rate in the Dutch population of fewer than 1 per 1,000 patients per year (Lamberts, 1984). Most 4th-year students in our medical school will have studied this disease in combination with other diseases that give pain in the upper abdomen such as gallbladder disease and other pancreatic diseases. Presumably, they will have paid special attention to its clinical appearance and to the way in which it can be differentiated from other upper abdominal diseases, which can be very difficult, because the signs and symptoms and lab findings of pancreatitis patients vary widely. Furthermore, as yet, its pathophysiology is not thoroughly understood. Hence, it is unlikely that 4th-year students will demonstrate a deep understanding of the condition. The relation between biomedical and clinical knowledge may be more prominent in other diseases (e.g., cholecystitis) and other medical specialties such as cardiology. Therefore, the observed delay in integration of biomedical and clinical knowledge may be specific for the kind of disease involved. It might be possible that, in domains with a tighter relation between biomedical and clinical knowledge, a smaller, or even no, gap will be found.

The curriculum format of medical schools also might affect the developmental course taken by medical students. Our subjects were from a school with a problem-based curriculum where biomedical sciences are taught in the context of specific cases posed as clinical problems. This approach to medical education has been developed in order to bridge the gap between the preclinical and the clinical period, a gap that requires students to relearn a great deal of the subjects they have studied and mastered before. Such a problem-based curriculum is thought to result in a better integration of biomedical and clinical knowledge. Research outcomes suggest that this might indeed be the case (Boshuizen & Schmidt, 1991). The study here, however, suggests that the gap still exists, and that students drastically change their approach to clinical problems. Whether their change is more or less drastic than the change students in a conventional curriculum experience is unknown. If, however, problem-based learning has the effects intended, then the gap experienced by students in a traditional school may be even greater.

Finally, can our conclusions be generalized to other, nonmedical domains of expertise? The answer to this question depends on two different elements: the educational approach and the task of the expert. The educational approach taken in medicine is in no way unique. Education in the professions at the college or university level almost always consists of a theoretical and a practical component. The theoretical part, in which basic concepts and sciences are taught, is scheduled first, followed by the practical part in which students are supposed to apply and practice what they have learned. This format is used for a wide variety of professions ranging from architecture to medicine and from engineering to business administration. Yet, the generic tasks of the professionals educated in this way may be very unlike one another. For example, a physician's task is diagnosis and remedying in order to heal patients, whereas architects have a very different cognitive task consisting of design and planning (Breuker et al., 1987).[8] An auditor's

[8] Breuker et al. (1987) developed a general model of expert problem solving. They assumed that domains may contain different concepts, but may share ways of using knowledge at some higher level of abstraction. This means that expertise in different domains may have the same higher level structure. Their taxonomy of generic tasks consists of three main classes: analysis, modification, and synthesis, further subdivided into about 20 generic tasks. The task structure in a specific domain of expertise can be described at an abstract level with one or more of these generic tasks.

job, on the other hand, is essentially the same as the physician's: Diagnosis (of a business firm's financial statements) and remedy. As the same generic task requires the same kind of knowledge structures, expert knowledge structures in medicine and auditing presumably will be more similar than in medicine and architecture.

Besides generic task similarities and differences, specific domain characteristics also affect the generality of our findings. One example is the amount of time spent on one problem. The average Dutch family physician sees about 25 to 30 patients a day. On the other hand, it takes an accountant about 2 to 3 weeks to audit the financial statements of a medium-size firm with 50 to 250 employees. Another example is the variation in the task due to external influences. In medicine this variation is small. Although medical equipment and methods of treatment change over time, the diseases themselves do not change dramatically in a specific population (despite exceptions like AIDS). Financial management and stockbroking on the other hand are very sensitive to variations in the economic situation. In case of a sudden economic decline (or a boom) the expertise built up before may become immediately useless. These examples suggest that the integration of theoretical and experiential knowledge may have different courses, varying in speed and smoothness, in the different domains of expertise.

## CONCLUSION

The two experiments presented here replicate the findings reported earlier by Patel and colleagues (Joseph & Patel, 1990; Kaufman & Patel, 1988; Patel, Arocha, & Groen, 1986; Patel, Evans, & Chawla, 1986; Patel & Groen, 1986; Patel et al., 1988). They provide a substantial basis for the conclusion that medical experts, contrary to intermediates and novices, do not overtly apply biomedical knowledge in clinical reasoning. In addition, our results expand the conclusions reached by Patel and colleagues in that we could show that expert biomedical knowledge is encapsulated and integrated into clinical knowledge. This latter finding contradicts the convictions of Patel et al. (1989) that biomedical and clinical knowledge essentially "represent two different worlds" and that, at least in routine cases, biomedical knowledge is not used at all. On the contrary, our findings suggest a tacit role of biomedical knowledge in expert clinical reasoning.

## REFERENCES


Ackermann, W., & Barbichon, G. (1963). Conduites intellectuelles et activité technique [Intellectual behaviors and technical activity]. *Bulletin CERP, 12,* 1-16.

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

Anderson, J.R. (1985). *Cognitive psychology and its implications.* New York: Freeman.

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

Boshuizen, H.P.A., & Schmidt, H.G. (1991, April). *Integration of biomedical and clinical knowledge in medical students.* Paper presented at the Annual Conference of the American Educational Research Association, Chicago, IL. (Maastricht: University of Limburg; O&O report.)

Breuker, J., Wielinga, B., Van Someren, M., De Hoog, R., Schreiber, G., De Greef, P., Bredeweg, B., Wielemaker, J., Billault, J.-P., Davoodi, M., & Hayward, S. (1987). *Model-driven knowledge acquisition: Interpretation models* (Deliverable task A1, Esprit Project 1098). Amsterdam: University of Amsterdam.

Chi, M.T.H., Feltovich, P.J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5,* 121–152.

Collins, A. (1990, April). *Generalizing from situated knowledge to robust understanding.* Paper presented at the Annual Conference of the American Educational Research Association, Boston, MA.

Collins, A., Brown, J.S., & Newman, S.E. (1989). Cognitive apprenticeship: Teaching the craft of reading, writing and mathematics. In L.B. Resnick (Ed.), *Knowing, learning and instruction: Essays in honor of Robert Glaser.* Hillsdale, NJ: Erlbaum.

Elio, R., & Scharf, P.B. (1990). Modeling novice-to-expert shifts in problem-solving strategy and knowledge organization. *Cognitive Science, 14,* 579–639.

Ericsson, K.A., & Simon, H.A. (1984). *Protocol analysis: Verbal reports as data.* Cambridge, MA: MIT Press.

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.* Assen, Holland: Van Gorcum.

Feltovich, P.J., Johnson, P.E., Moller, J.H., & Swanson, D.B. (1984). LCS: The role and development of medical knowledge in diagnostic expertise. In W.J. Clancey, & E.H. Shortliffe (Eds.), *Readings in medical artificial intelligence: The first decade.* Cambridge, MA: Addison-Wesley.

Hassebrock, F., & Prietula, M.J. (1986, October). *A protocol-based coding scheme for the analysis of medical reasoning.* Paper presented at the Eighth Annual Meeting of the Society for Medical Decision Making, Chicago, IL.

Joseph, G-M., & Patel, V.L. (1990). Domain knowledge and hypothesis generation in diagnostic reasoning. *Journal of Medical Decision Making, 10,* 31–46.

Kaufman, D., & Patel, V.L. (1988, August). Interactive medical problem solving in the context of a clinical interview: The nature of expertise. In *Proceedings of the 10th Annual Meeting of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum.

Kuipers, B.J. (1984). Common-sense reasoning about causality: Deriving behavior from structure. *Artificial Intelligence, 24,* 169–203.

Kuipers, B., & Kassirer, J.P. (1984). Causal reasoning in medicine: Analysis of a protocol. *Cognitive Science, 8,* 363–385.

Laird, J.E., Rosenbloom, P.S., & Newell, A. (1986). Chunking in SOAR: The anatomy of a general learning mechanism. *Machine Learning, 1,* 11–46.

Lamberts, H. (1984). *Morbidity in general practice: Diagnosis-related information from the monitoring project.* Utrecht, The Netherlands: Huisartsenpers b.v.

Leinhardt, G. (1987, April). *Situated knowledge: An example from teaching.* Paper presented at the Annual Conference of the American Educational Research Association, Washington, DC.

Lemieux, M., & Bordage, G. (1986). Structuralisme et pedagogie medicale: Etude comparative des stratégies cognitives d'apprentis-médecins. [Structuralism and medical education: A comparative study of the cognitive stratégies of novice physicians]. *Recherches Sémiotiques, 6,* 143–179.

Lesgold, A.M. (1984). Acquiring expertise. In J.R. Anderson & S.M. Kosslyn (Eds.), *Tutorials in learning and memory: Essays in honor of Gordon Bower.* San Francisco, CA: Freeman.

Lesgold, A., Rubinson, H., Feltovich, P.J., Glaser, R., Klopfer, D., & Wang, Y. (1988). Expertise in a complex skill: Diagnosing X-ray pictures. In M.T.H. Chi, R. Glaser, & M. Farr (Eds.), *The nature of expertise.* Hillsdale, NJ: Erlbaum.

Patel, V.L., Arocha, J.F., & Groen, G.J. (1986). *Strategy selection and degree of expertise in medical reasoning* (CME report). Montreal, Quebec, Canada: McGill University.

Patel, V.L., Evans, D.A., & Chawla, A. (1986). *Predictive versus diagnostic reasoning in the application of biomedical knowledge.* In Proceedings of the Eighth Annual Conference of the Cognitive Science Society. Hillsdale, NJ: Erlbaum.

Patel, V.L., Evans, A.E., & Groen, G.J. (1989). Biomedical knowledge and clinical reasoning. In D.A. Evans & V.L. Patel (Eds.), *Cognitive science in medicine: Biomedical modeling.* Cambridge, MA: MIT Press.

Patel, V.L., & Groen, G.J. (1986). Knowledge-based solution strategies in medical reasoning. *Cognitive Science, 10,* 91–110.

Patel, V.L., Groen, G.J., & Scott, H.S. (1988). Biomedical knowledge in explanations of clinical problems by medical students. *Medical Examiner, 22*(5), 398–406.

Patil, R.S. (1983). *Role of causal relations in formulation and evaluation of composite hypotheses.* Paper presented at the IEEE Med Comp Conference, Columbus, OH.

Patil, R.S. (1986). Review of causal reasoning in medical diagnosis. In *Proceedings of the Tenth Annual Symposium on Computer Applications in Medical Care.* Piscataway, NJ: IEEE.

Pople, H.E. (1982). Heuristic methods for imposing structure on ill-structured problems: The structuring of medical diagnosis. In P. Szolovits (Eds.), *Artificial intelligence in medicine.* Boulder, CO: Westview Press.

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.* Hillsdale, NJ: Erlbaum.

Schmidt, H.G., & Boshuizen, H.P.A. (in press). Encapsulation of biomedical knowledge. In D.A. Evans & V.L. Patel (Eds.), *Advanced models of cognition for medical training and practice.* Berlin: Springer-Verlag.

Schmidt, H.G., Boshuizen, H.P.A., & Hobus, P.P.M. (1988). Transitory stages in the development of medical expertise: The "intermediate effect" in clinical case representation studies. In *Proceedings of the 10th Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum.

Szolovits, P., Patil, R.S., & Schwartz, W.B. (1988). Artificial intelligence in medicine. *Annals of Internal Medicine, 108,* 80–87.

## APPENDIX A

### Symptoms and Pathophysiological Process in Pancreatitis

#### *Major Causal Factors, Cholelithiasis or Alcohol Abuse*

***Pathophysiological Process.*** Inflammation of the pancreas. The pancreas is an organ lying deep in the upper part of the abdomen, behind the peritoneum. Its function is to produce enzymes (e.g., lipase, amylase) that play an important role in the digestion of fat and protein. By their very nature, these enzymes are capable of destroying the pancreatic tissue as well. However, because they are released as inactive proenzymes, no harm is done. Furthermore, the pancreas produces insulin. Inflammation of the pancreas causes a swelling of the tissue, which may obstruct the pancreatic ducts. By this inflammation pancreatic proenzymes are activated in the gland itself, causing destruction of tissue. Pancreatic enzymes are released by this process, directly into the blood stream instead of the alimentary canal.

*Symptoms.* Pain in the upper part of the abdomen is often very severe, but, because of the location of the organ, physical examination may reveal only minor abdominal findings. Sometimes, pancreatitis is associated with malabsorption, caused by a lack of pancreatic enzymes in the alimentary canal. Because amylase is directly released into the blood stream, laboratory findings often show a high level of serum amylase.

## APPENDIX B

**Case Items**

1. Man, married, 38 years old, 2 children.
2. Vocation: gasfitter. Incapacitated (because of mental problems) last 4 years; does some odd jobs once in a while for his former employer.
3. **Prior illnesses.** Recurring neurotic depressions since the age of 24; has been hospitalized in a mental institution for 3 months; psychiatric outpatient since discharge.
4. **Prior illnesses.** Increasing alcohol abuse for 6 years.
5. **Prior illnesses.** Hospitalized 1 year ago because of an attack of acute pancreatitis.
6. **Prior illnesses.** Refusal® treatment 1 year ago.
7. Type of consultation: house call, 9:00 p.m.
8. Complaint: Continuous pain in the upper part of the abdomen, radiating to the back.
9. **History.** Started this morning with a vague pain in the abdomen.
10. **History.** Pain has increased since onset; constant now.
11. **History.** It is a heavy, boring pain, piercing right through.
12. **History.** Cannot localize the pain with one finger.
13. **Observation.** Attitude: The patient sits in a chair, continually bending over or pulling up his legs; otherwise moves normally.
14. **History.** Has just vomited; did not affect the pain.
15. **History.** The pain is similar to the pain he has been hospitalized with, but then it started very suddenly, and moreover it was worse.
16. **History.** He has had this sort of pain radiating to the back more often recently, but never as bad as it is now.
17. **History.** Most of the time it vanished within a day.
18. **History.** He tried some Aspirins® but they didn't have any effect.
19. **History.** Alcohol abuse: Drinking again for past 6 months; drinks the same amount as he used to.
20. **History.** Depressions: Has been feeling low for last 10 days; 2 days ago he tried to put an end to everything with 6 Mogadon and quite a lot of alcohol.

21. **Observation.** Appearance: Doesn't look ill, but his face is distorted with pain.
22. **Observation.** Appearance: Not anemic, not jaundiced.
23. **History.** According to other people, he has been jaundiced once in a while recently.
24. **History.** Appetite: Poor.
25. **History.** Eating habits: Irregular.
26. **History.** Eating habits: Often eats chips and a salad, a sausage or a meat ball.
27. **History.** Today's menu: He has had breakfast with ham and eggs and bread. Dinner yesterday: spaghetti with tomato sauce.
28. **History.** Pain has no relation to meals.
29. **History.** Weight: Has lost ±5 kg in the last 6 months.
30. **History.** Defecation: No problems, defecation pattern not changed.
31. **History.** Defecation: Paler and more malodorous stools according to the patient.
32. **History.** Defecation: Last bowel movement was yesterday.
33. **History.** Temperature: 37.8 °C at 6 p.m.
34. **Physical examination.** Pulse rate: regular, 72/min.
35. **Physical examination.** BP: 140/95.
36. **Physical examination.** Respiration rate: 18/min.
37. **Physical examination.** Abdomen, Inspection: Moderately distended, moderately moving with respiration.
38. **Physical examination.** Abdomen, auscultation: Diminished bowel sounds.
39. **Physical examination.** Abdomen, percussion: Liver and spleen not enlarged.
40. **Physical examination.** Abdomen, palpation: Epigastric tenderness; no defence; no further tenderness in the abdomen; no further palpable anomalies in the abdomen.
41. **Physical examination.** Cor, Percussion: not enlarged; normal sound; regular rate; no murmurs.
42. **Physical examination.** Pulmones: Both boundaries normally moving; normal respiration sounds.
43. **Physical examination.** Arteria femorales, palpation: Good pulsations at both sides; auscultation, no abnormal murmurs.
44. **Physical examination.** Rectal examination: No anomalies.
45. **Physical examination.** Venous pressure not raised.
46. **Physical examination.** Parotic glands not enlarged.
47. **History.** No further complaints.
48. **Lab findings:** Serum amylase: 128 U (normal 8-16). Glucose: 6.0 mmol/l (normal 4.4-5.8). WBC: $11.0 \times 10^9$/l (normal $5-10 \times 10^9$). ESR: 15 mm after 1 hour (normal – 10). Hg: 7.8 mmol/l (normal 8.8-11.2).

## APPENDIX C

Responses of Subject I-1 to 16 Items and the Propositions Extracted

| Item Number | Subject's Response | Propositions Extracted |
|---|---|---|
| 8. Complaint: continuous pain in the upper part of the abdomen, radiating to the back. | ...may have to do with the *gall bladder.*\|\|may have to do with the *pancreas*...\|\| | #8→(H) gall bladder →(H) pancreas |
| 9. **History.** Started this morning with a vague pain in the abdomen. | I couldn't directly...make anything of that\|\| | |
| 10. **History.** Pain has increased since onset; constant now. | ..so that's against eh biliary eh colic (I) em\|\|. so *gallstones*\|\|....but further?\|\| | #10→(cond neg) (I) biliary colic →(cond neg) (H) gallstones |
| 11. **History.** It is a heavy, boring pain, piercing right through. | ...I can't make anything of that.\|\| | |
| 12. **History.** Cannot localize the pain with one finger. | ......that means that it's a rather diffuse.eh diffuse pain...\|\|.only not complete(ly), yes. eeehm.\|\|. | #12→(I) rather diffuse pain |
| 13. **Observation.** Attitude: The patient sits in a chair, continually bending over or pulling up his legs; otherwise moves normally. | ...(he eh pulls up his legs against his body), which means that he tries to avoid contraction of his abdominal muscles\|\|that means that probably something is wrong with the *peritoneum,*\|\|a *peritonitis* possibly a beginning one.\|\| and that might be an indication of a..*perforated peptic ulcer* (H)..\|\| | #13→(I)* avoid contraction of his muscles →(cond) (H) something wrong with the peritoneum →(spec) (H) beginning peritonitis →(cond) (H) perforated peptic ulcer |
| 14. **History.** Has just vomited; did not affect the pain. | ..that may be the case when a eh, when the *peritoneum* is involved too\|\| | #14←(cau) (H) peritoneum involved |
| 15. **History.** The pain is similar to the pain he has been hispitalized with; but then it started very suddenly, and moreover it was worse. | ....ehm.now I don't know what his...\|\| (yes yes a year ago it was acute pancreatitis)\|\|.ehm... doesn't say anything in particular;\|\|only still points in the direction of eh..eh that *peritonitis eh irritation*\|\| | #15→(cond) (H) peritonitis irritation |

*Note.* The subject's responses are subdivided into protocol fragments by means of ||; repetitions of findings are placed in parentheses.
H=hypothesis, I=interpretation, *=biomedical.