Memory & Cognition
1993, 21 (3), 338-351

On the origin of intermediate effects in clinical case recall

HENK G. SCHMIDT and HENNY P. A. BOSHUIZEN
University of Limburg, Maastricht, The Netherlands

In two experiments, the effects of level of medical expertise and study time on free recall of a clinical case were assessed. In Experiment 1, a nonmonotonic relationship between level of expertise and recall was found: Subjects of intermediate levels of expertise remembered more information from the case than both experts and novices. This "intermediate effect" disappeared, however, when study time was restricted. Analysis of post hoc acquired protocols of pathophysiological knowledge active during case processing suggested that this phenomenon could be attributed to the nature of the pathophysiological knowledge mobilized to comprehend the case. In Experiment 2, this assumption was directly tested by priming relevant pathophysiological knowledge for either a short or a longer period, before enabling subjects to study the case briefly. Free-recall data confirmed and extended the results of Experiment 1. Again, an intermediate effect was found; this time, however, it was generated experimentally. The findings were interpreted in terms of qualitative differences in the nature of the knowledge structures underlying performance between novices, advanced students, and medical experts: Experts use knowledge in an encapsulated mode while comprehending a case, whereas students use elaborated knowledge.

The *intermediate effect* in clinical case representation studies is among the best-known, stable, and hitherto unexplained phenomena in medical expertise research. The quasi-experimental paradigm that produces this phenomenon is described as follows: Subjects differing in level of expertise are requested to study, for about 2 or 3 min, half a page of text describing a patient's history, presenting complaint and some additional findings such as results of laboratory tests and physical examination. The text is removed, and the subjects are asked to recall everything they can remember from the text. Subjects of intermediate levels of expertise consistently produce more elaborate recalls than either experts (e.g., experienced physicians) or novices. This phenomenon has been demonstrated under various conditions, with different cases and in different populations (Claessen & Boshuizen, 1985; Hassebrock, Bullemer, & Johnson, 1988; Muzzin, Norman, Feightner, & Tugwell, 1983; Patel & Groen, 1986b). The intermediate effect has also been demonstrated in expertise-related tasks other than text processing (Grant & Marsden, 1988; Patel, Evans, & Kaufman, 1988).

These findings appear to be counterintuitive. Spilich, Vesonder, Chiesi, and Voss (1979), for instance, have shown that subjects with a high knowledge of baseball remembered more, and more relevant, information from a report of a baseball game than low-knowledge individuals. Theories of text processing, generally, assume that prior knowledge facilitates new information to be encoded and retrieved. The more prior knowledge a person has, the more he/she will be able to recall from the stimulus material (Graesser & Clark, 1985; Voss & Bisanz, 1985). In other words, instead of the inverted U-shaped curve commonly found in developmental studies in medicine, one would expect a monotonically increasing recall function with increasing expertise.

In the present article, two experiments that were conducted to find an explanation for this apparent paradox will be reported. First however, clinical case representation studies relevant to the present perspective will be briefly reviewed.

Representation of Clinical Cases by Medical Students and Experts

Studies using free recall of meaningful texts assume that the resulting data reflect the contents and the structure of the mental representation of the original stimulus material. The mental representation itself emerges from the interaction of the stimulus material (e.g., a description of a clinical case) and prior knowledge relevant to that material. The input text activates prior knowledge, which in turn provides meaning and structure to the incoming information (Kintsch, 1988; Van Dijk & Kintsch, 1983). Kintsch and Greeno (1985) have introduced the term *problem model* to describe the mental representation originating from an attempt to comprehend and solve a problem. It is assumed that the better the understanding of the per-

Parts of this article were presented at the August 1988 meeting of the Cognitive Science Society in Montreal, Canada. Preparation was enabled by Grant 6626 from the Institute for Educational Research in The Netherlands (SVO) to H. G. Schmidt and H. P. A. Boshuizen and by a grant from the Spencer Foundation to H. P. A. Boshuizen. Requests for reprints should be sent to H. Schmidt, Department of Educational Development and Research, University of Limburg, P.O. Box 616, 6200 MD Maastricht, The Netherlands.

son attempting to solve a particular problem, the more elaborate and adequate the problem model will be. Since experts have a far more elaborate knowledge base than novices or intermediates do, one would expect their problem model, and hence their free recall of the problem, to be more elaborate. This is essentially what has been found in a number of domains, such as computer programming (McKeithen, Reitman, Rueter, & Hirtle, 1981; Shneiderman, 1976), go (Reitman, 1976), football results (Morris, Gruneberg, Sykes, & Merrick, 1981), mental calculation (Hunter, 1962), and chess (Chase & Simon, 1973; de Groot, 1946). For instance, de Groot (1946) presented advanced chess players, masters and grand masters, with midgame positions and required them to recall the positions of the pieces on the board. He demonstrated that accuracy and amount of recall was positively related to level of expertise.

The first free-recall studies in medicine were conducted by Norman and his associates (Muzzin et al., 1983; Norman, Jacoby, Feightner, & Campbell, 1979), who presented medical students and family physicians with short written cases. The subjects were requested to study each of these cases, to recall the information presented, and to produce a diagnosis. In one study, in which the subjects were requested to process four cardiorespiratory cases, each within 45 sec, it was found that students in residency training and third-year medical students showed more recall than the physicians (Muzzin et al., 1983). Claessen and Boshuizen (1985) presented cases of pancreatitis and prostatitis to family physicians and three groups of students at different levels of expertise. Study time was free. The residents participating in the experiment showed superior recall on all cases as compared with both the family physicians and the preclinical students. Patel and Medley-Mark (1985) demonstrated the same phenomenon in final-year medical students as compared with both internists and novices, using an acute bacterial endocarditis and a stomach cancer case, each studied for about 2.5 min. Boshuizen (1989) presented two cases to four groups of subjects—second-, fourth-, and fifth-year medical students and family physicians. The cases were presented in a sequential fashion and the subjects were encouraged to think aloud while studying the patient information. Study time was free. For each of the cases, she found a significant quadratic relationship between level of expertise and free recall; the subjects of intermediate levels recalled more information than both family physicians and second-year students. Hassebrock et al. (1988) presented advanced students, residents, and pediatric cardiologists with two cases in pediatric cardiology, not constraining study time. They failed to find differences in the immediate recall of the cases, however, after a delay of 1 week, the cardiologists remembered significantly less. Patel and Groen (1991) reviewed several studies involving the comprehension of clinical cases by subjects of different levels of expertise. In each of these studies they found that intermediates recalled more information than both novices and experts. Two other studies failed to show significant differences in the recall of clinical cases (Muzzin et al., 1982; Norman et al., 1979).

Thus, in contrast to studies from domains such as chess or computer programming, clinical case recall studies either demonstrate more elaborate recall by subjects of intermediate levels of expertise, or fail to find significant differences. Not only do these outcomes appear to be at variance with general models of text processing (e.g., Graesser & Clark, 1985), they also contradict existing theories of expertise in medicine. Lesgold and his associates (Lesgold, 1984; Lesgold et al., 1988), for instance, emphasize the determining role of causal, biomedical knowledge in expert performance. They assume that experts in the medical domain have acquired a more elaborate biomedical knowledge base, enabling them to perform better than subexperts on diagnostic and other expertise-related tasks, such as recognizing malignant structures on X-rays (see also Elstein, Shulman, & Sprafka, 1978). It is difficult, however, to see how this conception of expertise in medicine could explain a *decrease* in recall of clinical cases associated with an extension of the knowledge structures assumed to be responsible for the recall performance.

A recent reformulation of Van Dijk and Kintsch's (1983) theory of text comprehension by Kintsch (1988) may elucidate in more detail why current theories of text processing have difficulty explaining this anomaly. Kintsch's construction-integration model deviates from existing schema-based conceptions of text comprehension in that it conceptualizes the processing of text as largely bottom-up. It assumes that knowledge is not prestored in fixed structures, but is generated in the context of the task for which it is needed. Knowledge about the world can be conceived as an associative network of nodes and their interrelations. When a reader attempts to understand a text, words or phrases will activate their corresponding nodes and activation will spread to other related nodes. In this way, a pool of knowledge is activated that may or may not be relevant to the task of understanding the text at hand. Kintsch describes this as the "construction" phase in text comprehension. Further reading, and hence activation, of concepts and their interrelations will, however, constrain the meanings of what is previously read by deactivating irrelevant knowledge. Thus, a coherent representation of the text is formed, integrating knowledge from earlier cycles of activation with that of later cycles. In the course of this integration process, the reader may have to produce bridging inferences and form macropropositions to create or maintain coherence. Prior knowledge and information from the text itself thus become integrated into a text base that represents the meaning of that text. Extending Kintsch's argument, it seems to be a reasonable assumption that the more prior knowledge a person has, the less time is needed to construct a coherent text base.[1] Experts can be expected to produce less bridging inferences, simply because these are already part of their knowledge base and will be generated automatically. This would explain why experts are generally faster and more accurate in text-processing tasks related to expertise (e.g., Chi, Glaser, & Farr, 1988; Spilich et al., 1979). Because expertise transfers the advantage of speed to processing, experts can be expected to produce more elaborate recalls than subjects of less advanced levels of expertise. The notions proposed by Kintsch thus leave unexplained why, compared with intermediates, experts in medicine, in fact, produce *less* elaborate case recalls.

In an attempt to fit these findings to the Van Dijk and Kintsch (1983) model, Groen and Patel (1988) assume that subjects reading a text about a patient transform the discourse into two different kinds of representations—the textbase, which is the semantic representation of the input, and a situation model, which is the cognitive representation of the events, actions, persons, or the situation in general that the text is about. Groen and Patel suggest that experts have better situation models than intermediates do, enabling them to "filter out" irrelevant information from a case. This filtering out would be responsible for the fact that experts recall less information; their situation models induce them to pay attention only to cues that are critical to the situation and discard others. However, the issue here is not so much *that* experts may filter out irrelevant information, but *why* they are able to do so. That is, what kinds of changes occur in the underlying knowledge base so that the representation of a clinical case becomes more condensed when expertise develops beyond the intermediate level?

Schmidt and Boshuizen (1992, in press; Boshuizen & Schmidt, 1992) have proposed the idea of *knowledge encapsulation* to explain this phenomenon. Knowledge encapsulation is the subsumption, or "packaging," of lower level detailed propositions, concepts, and their interrelations in an associative net under a smaller number of higher level propositions with the same explanatory power. They suggest that novices and intermediates process a clinical case much in the way described by Kintsch (1988), using increasingly detailed and elaborate biomedical knowledge about the condition of the patient described. This process is essentially bottom-up, so these groups conduct extensive search while interpreting the signs and symptoms displayed in the story, leading to increases in recall as a function of the growth of the knowledge base. However, as more and more similar patients are seen, certain shortcuts in their thinking begin to emerge. Sets of detailed propositions originally activated while comprehending a case become encapsulated in concepts of greater generality. Under the influence of experience, subjects' elaborate knowledge of the world becomes "chunked" into a limited number of highly inclusive concepts that have the same explanatory power as the original elaborate structure. In the case of medicine, the encapsulating concepts often tend to be of direct clinical relevance.

Knowledge encapsulation was first observed in "think-aloud" and pathophysiological explanation protocols of medical experts. When asked to think aloud about a clinical case, or to explain the phenomena observed in a case in terms of their underlying structure, experts appear to apply less biomedical knowledge than either intermediates or even novices. Boshuizen, Schmidt, and Coughlin (1987), for instance, showed that general practitioners rarely refer to biomedical concepts while reasoning about a case, whereas students use biomedical concepts extensively in the same situation. Patel, Evans, and Groen (1989), who reviewed research on pathophysiological explanations of clinical cases by subjects of different expertise, also conclude that experts appear to rely less on causal biomedical knowledge while diagnosing a case, and tend to explain cases largely in terms of clinical knowledge.[2] Boshuizen and Schmidt (1992) have demonstrated that biomedical knowledge is largely missing from experts think-aloud protocols, not so much because this kind of knowledge has become inert or inaccessible, but because, through practice, biomedical concepts have become encapsulated into higher level concepts. Their findings suggest that, as a result of the extended application of knowledge in the comprehension of clinical cases, the granularity of the concepts applied changes, with experts using more global, comprehensive concepts than subjects of lower levels of expertise. Hence, the development of expertise seems to involve increased processing speed and further expansion of the knowledge base as well as qualitative shifts in the knowledge base itself.[3]

If this analysis is correct, what are the implications for case recall by experts? We hypothesize that encapsulation of concepts in the knowledge base is directly associated with encapsulation of related information in a text. Experts perceive and process individual signs and symptoms as integrated wholes (for which they have encapsulating concepts) and recall these encapsulating concepts rather than the individual items themselves. For instance, if one would require an internist to recall a patient history of a toxic-looking drug addict who presents himself with high fever, shaking chills, sweating, and feelings of prostration with some shortness of breath and a high pulse rate, the internist would tend to process, and hence to recall, this set of items as "patient has a septic condition." "Septic condition" is an inference from the text that summarizes or encapsulates these individual items. This would explain why experts' recall is sparser than intermediates' recall of a clinical case: Experts recall the case in encapsulated mode. In addition, it would explain why experts produce more inferences in recall, although their overall recall is less (Coughlin & Patel, 1986; Patel & Medley-Mark, 1985). If we are correct, these inferences must be encapsulations of the original text.

This theory explaining the origin of differences in problem representation between subjects of various levels of expertise allows for a number of predictions that were tested in Experiment 1:

1. If the construction and integration of biomedical knowledge to understand the information embedded in a case is more extensive than simply matching relevant knowledge in encapsulated mode, intermediates may be

expected to recall the case more extensively than experts. This would explain the existence of intermediate effects.

2. The search for relevant concepts and the production of bridging inferences to form a coherent representation, however, takes more time than activating a relatively small number of comprehensive, encapsulating concepts. Hence, it is predicted that the intermediate effect will disappear when processing time is restricted. Under these conditions, students will be less able to form a coherent text base. Experts encode sets of individual items in encapsulated mode without much search, so their recall will be less affected by a reduction of processing time.

3. If experts represent case information in encapsulated form, their recall will contain more comprehensive, high-level inferences.

4. If intermediates usually process a case by activating elaborate biomedical knowledge, whereas experts only apply encapsulated knowledge, then post hoc pathophysiological explanations provided by intermediates will be more extensive than those provided by experts.

5. This will only apply to those circumstances in which subjects have sufficient time to use their pathophysiological knowledge. If the time available is critical to the search, activation, and processing of relevant knowledge, a decrease in processing time will also affect the elaborateness of pathophysiological explanations provided by intermediates. The amount of explanation provided by the experts, however, will not be influenced by manipulation of processing time, because their knowledge is encapsulated and easily available.

6. Because of encapsulation, pathophysiological explanations by experts will contain more comprehensive, high-level propositions than those of intermediates.

In Experiment 1, these predictions were tested by requiring subjects of different levels of expertise to study a case history under varying time constraints, recall the information, provide a diagnosis, and produce a pathophysiological explanation for the signs and symptoms described in the case.

EXPERIMENT 1

Method

Subjects. The subjects were 120 students and physicians of the University of Limburg—24 first-year allied health sciences students (the "laypersons" in the present study), 72 medical students (24 second-, 24 fourth- and 24 sixth-year undergraduate students),[4] and 24 internists with at least 2 years of experience. Each group of 24 was randomly subdivided into three groups of 8, studying the clinical case under different time constraints. The subjects received a small compensation for their participation.

Materials. The materials consisted of a booklet containing a 270-word description of a clinical case and three blank response sheets. The case was a Dutch translation of the acute bacterial endocarditis case used by Patel and Groen (1986a) and consisted of 71 propositions. The case is considered to be of intermediate difficulty, that is, not entirely routine even for experts. In the Patel and Groen study, for instance, six out of eight cardiologists produced an accurate diagnosis. The original text of the case is provided in Appendix A.

Procedure. The subjects were tested individually. First, they were requested to study the case carefully. Depending on the experimental condition, they were required to study the case for either 3 min 30 sec (3.5 min)—the amount of time allotted to the subjects in the original Patel and Groen study—1 min 15 sec (1.25 min), or 30 sec. Before processing the case, the subjects in the 30-sec condition were given the opportunity to read an unrelated text of exactly the same length to provide them with some experience in scanning a text in a very short time. This was done to minimize variability in the way they would undertake the experimental task. After reading the text for the duration of time allowed, the experimenter instructed them to turn to the next page, which contained the following instructions: "Please, write down everything you can recall from the case. Write complete sentences and avoid abbreviations." On the next page, the subjects were requested to provide a diagnosis for the patient. The last page contained these instructions: "Describe the pathophysiology which, in your opinion, underlies the case. Write complete sentences and avoid abbreviations. Use schematic representations only if strictly necessary." The subjects were free to use as much time as they needed for the assignments.

Analysis. The accuracy of the diagnosis was determined by attaching weights to each of the elements. If the diagnosis contained the term "endocarditis," 2 points were given. The presence of "acute," "bacterial," "emboli," or "contaminated needles" each contributed 1 point. So, a maximum score of 6 could be obtained.

Both recall and pathophysiological explanation protocols were segmented into propositions by using a technique adapted from Kintsch (1974). The propositions consisted of two concepts connected by a qualifier, such as "causation (cau)," "negation (neg)," "location (loc)," or "specification (spec)." For instance, the protocol fragment "(. . .) The ictus is not displaced and a 2/6 diastolic murmur is heard over the aortic valve (. . .)" consists of four propositions: 1. ictus—neg (displaced); 2. murmur heard—loc (over the aortic valve); 3. murmur—spec (2/6); and 4. murmur—spec (diastolic). For each proposition in the free recall, it was decided whether it matched any proposition in the stimulus material. Inferences were included in the analysis to the extent that they could be matched to one or a combination of propositions in the original text. Repetitions were removed. Interrater agreement for this procedure was 93%. The total numbers of propositions found in free recall and pathophysiology protocols were tallied.

In free recall, the evidence for encapsulation of signs and symptoms was explored by counting the number of high-level inferences (Coughlin & Patel, 1986), which were inferences that summarized at least two propositions from the text. This is a subset of the total number of inferences produced. Inferences referring to only one proposition in the text were excluded, because these could not be considered encapsulations. To distinguish the former from inferences based on only one proposition in the text, they will be further referred to as *summaries*.

To test the hypothesis that the physicians' pathophysiology protocols contained knowledge in encapsulated mode, that is, contained more comprehensive high-level propositions than those of the students, additional analyses were carried out. Each of the propositions in the pathophysiology protocol was matched against a canonical explanation, provided in Appendix B. The canonical model contains 11 highly encapsulating propositions, which constitute a minimally sufficient explanation for all the signs and symptoms in the endocarditis case.[5] It was constructed with the support of two internists familiar with the disease and its manifestations. Each of the 11 explanatory concepts mentioned in the canonical model represent knowledge at a high level of encapsulation. The number of propositions matching the model's propositions was counted. Interrater agreement was higher than 90% for each of these procedures.

The data were analyzed using multivariate analyses of variance. Polynomial contrast analyses were conducted to test for nonlinearity of the relations between the dependent variable and the subjects' levels of expertise. This analysis calculates linear, quadratic, cubic, and quartic terms in the between-groups variance. The predicted inverted U-shape in the recall data would result in significant deviations from linearity and a significant quadratic term without significant deviations.

Results

Diagnostic accuracy. Figure 1 shows the average accuracies of the diagnoses proposed by the subjects. The number of accurate diagnoses significantly covaried with expertise, resulting in a monotonically increasing performance curve [$F(4,105) = 25.52$, $MS_e = 1.46$, $p < .0001$]. These data suggest that the case representation task is ecologically valid, because the expected expertise-related differences were actually found. Processing time had an overall effect on diagnostic accuracy [$F(2,105) = 3.163$, $MS_e = 1.46$, $p < .05$]. None of the possible within-group comparisons were statistically significant, however, including expert performance [$F(2,21) = 2.62$, $MS_e = 3.88, p < .10$]. This may be due to a lack of statistical power, because each of the groups compared consisted of only 8 subjects. As the average scores of the student groups were low under all conditions, failure to find significant differences within levels of expertise there may be the result of the overall difficulty of the case.

Free recall. Figure 2 shows the results of the analyses of the free-recall protocols. The number of propositions recalled is displayed as a function of expertise and processing time. Overall differences between levels of expertise were statistically significant [$F(4,105) = 5.25$, $MS_e = 70.40$, $p < .001$]. The relation between expertise level and recall was nonlinear, with a significant quadratic component [$F(1,115) = 7.02$, $MS_e = 167.50, p < .01$]. In other words, an inverted U-shaped curve provides the best fit to the data. In addition, processing time had a significant effect on recall [$F(2,105) = 68.26$, $MS_e = 70.40$, $p < .0001$].

For the different processing conditions, the following patterns emerge: In the 3.5-min processing condition, effects of expertise level on recall were significant [$F(4,35) = 5.64$, $MS_e = 118.85, p < .001$], with both a significant linear and a quadratic component. Under the 1.25-min processing condition, differences among expertise levels were nonsignificant [$F(4,35) = 2.18$, $MS_e = 68.51$, $p < .10$]. However, the 30-sec processing condition again displayed a significant effect of expertise on case recall [$F(4,35) = 4.81$, $MS_e = 23.84, p < .01$], with a significant linear component.

Pairwise comparisons within conditions show the following results: Under the 3.5-min processing time condition, the three intermediate levels of expertise produced significantly more propositions than the internists and the health sciences students, nicely illustrating the existence of an intermediate effect in these data. Among the groups that were required to process the case in 1.25 min, only the fourth- and sixth-year students produced significantly more recall, indicating that the intermediate effect, although still traceable, is less prominent. In the shortest processing condition, the two highest levels of expertise differed significantly from the other groups, but not from each other.

In summary, as predicted, the emergence of the intermediate effect appears to be dependent on the amount of time available for processing clinical information. Given a sufficiently short period of time, the phenomenon disappears and the well-known increase in recall performance as a function of expertise materializes (Chase & Simon, 1973; de Groot, 1946).[6] It is interesting to note that differences *within* levels of expertise as a result of differences in processing time were highly significant (all *p*s < .0001), with the exception of the internists [$F(2,21) = 1.34$, $MS_e = 80.43$, $p = .28$]. These data suggest that, in contrast to students of all levels of expertise, the performance of experienced physicians is relatively insensitive to manipulations of time, at least within the limits of the present experiment.

Figure 3 displays the number of summaries found in the free-recall protocols. Overall differences between

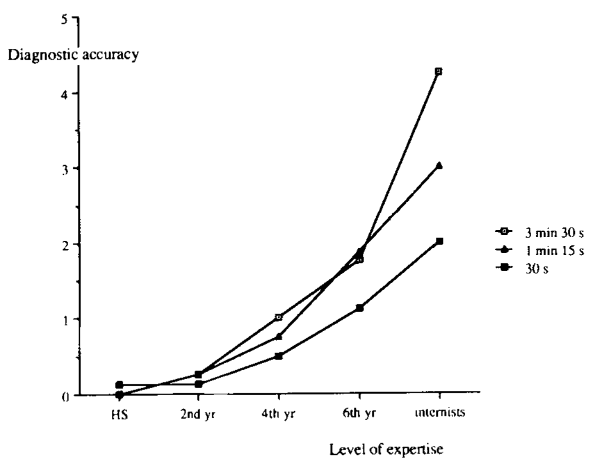


Figure 1. Average accuracy of diagnoses as a function of expertise and processing time.

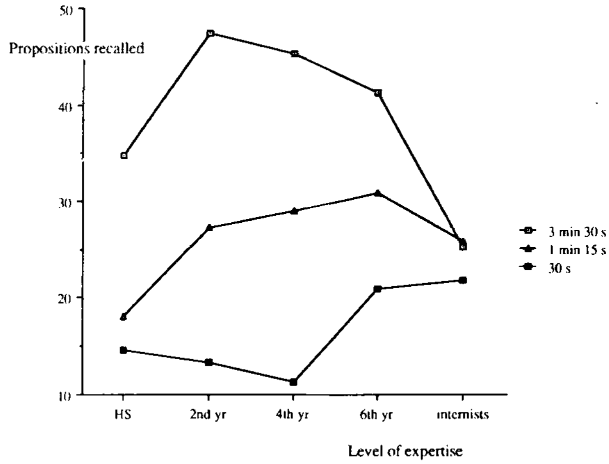


Figure 2. Number of propositions recalled from the acute bacterial endocarditis case as a function of expertise and processing time.

levels of expertise were again statistically significant [$F(4,105) = 25.862$, $MS_e = .813$, $p < .0001$]. Both a linear and a quadratic model can adequately describe the data [$F(1,115) = 88.16$, $MS_e = .72$, $p < .0001$, and $F(1,115) = 20.99$, $MS_e = .72$, $p < .0001$, respectively]. Processing time had no significant effect on number of summaries produced [$F(2,105) = .63$, $MS_e = .81$, $p < .54$]. Summaries were defined as inferences encapsulating two or more propositions appearing in the text, so these data suggest that experts recall more information from a text in encapsulated mode than subjects of lower levels of expertise.

Pathophysiological explanation protocols. The assumption was that intermediates and experts use functionally different knowledge while representing a clinical case. According to Schmidt and Boshuizen (in press), intermediates process information by activating a rich base of detailed causal pathophysiological knowledge, whereas experts activate encapsulated knowledge. Assuming that post hoc pathophysiological explanations reflect the knowledge activated during the case processing task (Groen & Patel, 1988), the intermediates' protocols will be more extensive than those provided by the experts. However, if the intermediates are hindered in activating their knowledge by restricting the time needed to activate relevant pathophysiological knowledge, information from the case cannot properly be processed. Under the latter condition, the elaborateness of post hoc explanation by intermediates is expected to decrease, whereas pathophysiological explanations provided by the experts—originating from encapsulated knowledge and therefore less extensive—will remain fixed. Figure 4 contains quantitative information concerning this assumption.

Both effects of level of expertise and processing time were statistically significant [$F(4,105) = 6.49$, $MS_e = 74.02$, $p < .0001$, and $F(2,105) = 11.15$, $MS_e = 74.02$, $p < .0001$, respectively]. The effect of expertise level had a significant quadratic component [$F(1,115) = 19.03$, $MS_e = 89.01$, $p < .0001$], indicating that subjects of intermediate levels produce more elaborate pathophysiological explanations than both novices and experts. Overall differences were highly reliable, both within levels of the experimental treatment and within each of the levels of expertise. Exceptions were the sixth-year students [$F(2,21) = 2.28$, $MS_e = 142.96$, $p < .13$] and the internists [$F(2,21) = .40$, $MS_e = 36.40$, $p < .67$]. Among students, constraining processing time generally caused a decrease in the number of propositions produced in the pathophysiology protocols. Because the subjects were free to use as much time as they needed to produce the protocols, these results strongly suggest that less extensive activation of pathophysiological knowledge took place in these groups while trying to understand the clinical case.

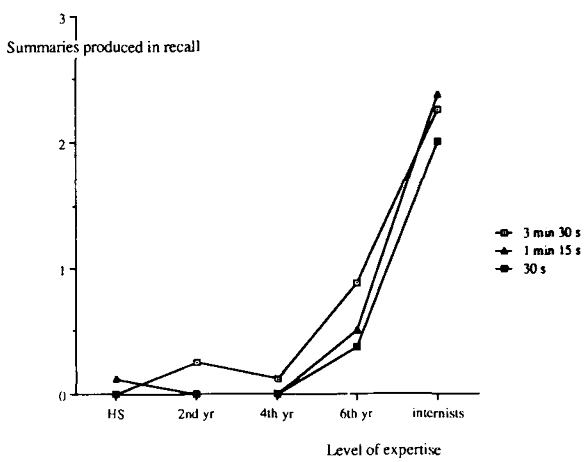


Figure 3. Number of high-level summaries produced as a function of expertise and processing time.

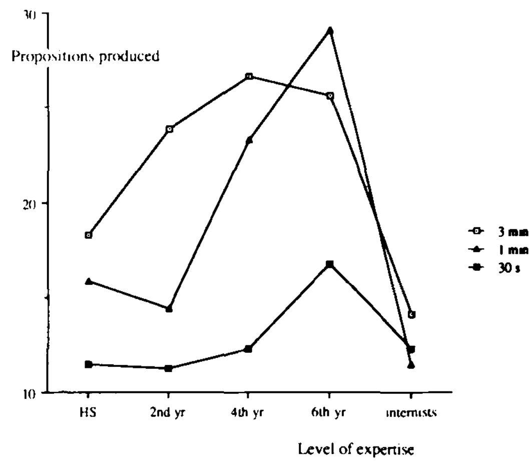


Figure 4. Number of propositions found in the pathophysiology protocols as a function of expertise and processing time.

In many ways, the pathophysiology data are similar to the free-recall data. An intermediate effect is present, and processing time seems to have the same effect on pathophysiological reasoning as it has on free recall. These similarities indeed suggest that the intermediate effect in free recall is the result of extensive activation and processing of pathophysiological knowledge in the course of understanding the case and that failure to do so results in poor recall of the information embedded in that case. Again, the physicians appear to be a separate case. Their output was small and quite stable, irrespective of the constraints on processing time.

Finally, it was predicted that the physicians' pathophysiology protocols would contain more propositions in encapsulated mode than the students' groups, which would explain why their explanations are more condensed than those of students. To test this hypothesis, the contents of these protocols were matched against the canonical model described in Appendix B, which contains highly encapsulating explanatory propositions. If encapsulation is indeed a function of expertise, a linear trend is to be expected in the data. Figure 5 displays the average numbers of matches between the canonical propositions and the propositions found in the protocols.

Both expertise level and processing time were statistically significant [$F(4,105) = 21.03$, $MS_e = 2.26$, $p <$

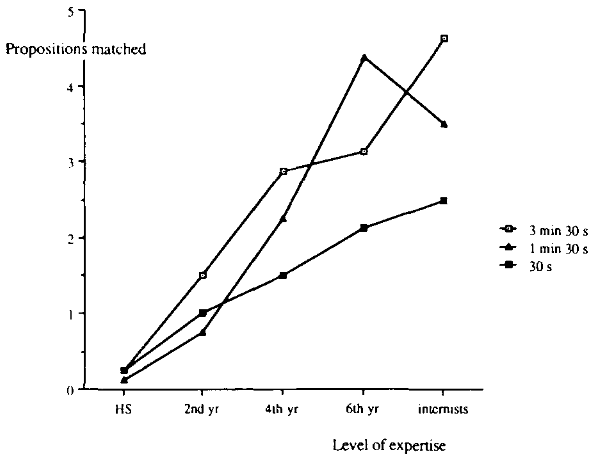


Figure 5. Number of encapsulating concepts found in the pathophysiology protocols as a function of expertise and processing time.

.0001 and $F(2,105) = 4.72$, $MS_e = 2.26$, $p < .01$, respectively]. No higher order trend was found, indicating that the number of encapsulating propositions produced linearly increased with expertise. (Hence, the better performance of the sixth-year students in the 1.5-min condition must be attributed to chance.) In summary, although the total number of pathophysiological propositions produced by the experts was much lower than that produced by the intermediates (as exemplified in Figure 4), the number of encapsulating propositions was higher. These data suggest that the use of knowledge in encapsulated form is indeed a function of expertise.

Discussion

The data presented support the notion proposed by Schmidt and Boshuizen (in press) that students and experienced physicians represent clinical cases in different ways because, in the process of understanding the text, the groups use functionally different knowledge. Provided they have sufficient time, medical students seem to elaborately process causal pathophysiological knowledge activated by cues embedded in the text in a bottom-up fashion, trying to construct a more or less coherent text base (Kintsch, 1988). By contrast, experienced physicians activate only a few concepts relevant to understanding the text. This would explain why advanced students recall more information from a case than both experts and novices, which accounts for the emergence of an intermediate effect. This processing explanation of the intermediate effect was further supported by the effects of time constraints. As predicted, the intermediate effect disappeared when less time was available for processing. This experimental manipulation greatly affected the performance of novices and intermediates, whereas the experts' recall was largely unaffected. Since activating relevant concepts in encapsulated mode takes less time than construction and integration of a text base through activation of detailed knowledge and extensive production of bridging inferences, the effects of time constraints on intermediates, but not on experts, is easily interpreted. Finally, although the experts' overall recall was less (in the 3.5-min condition it was even less than the laypersons'), they produced five times more summaries in their recall than even the sixth-year medical students, suggesting that they process the information in an encapsulated form. One internist, for instance, recalled the text as follows: "The patient is a young man with a high fever, who presents a *septic syndrome*. This suggests *drug use*. He shows signs of *thromboemboli*, due to an *affected heart valve*. The *tachycardia* fits with an *associated aorta vitium* [italics added]." His protocol consists almost entirely of summaries of the information actually presented in the text. These summaries were defined as high-level inferences, condensing many individual items from the text. For instance, an inference such as "septic condition" in itself subsumes 26 propositions from the endocarditis text. The same applies to inferences such as "hemodynamic stability" and "aorta malfunction," often showing up in the recall protocols. It is interesting to note that the same terminology also often appears in the pathophysiological explanations, indicating that experts tend to perceive, and hence describe, the world in terms of its underlying structure rather than in terms of its surface characteristics, a finding also reported by other investigators of expertise development (Chi, Feltovich, & Glaser, 1981; Chi et al., 1988).

The post hoc pathophysiology protocols were in line with the predictions. If the subjects had had only limited time to study the case, their post hoc pathophysiological explanations of the signs and symptoms encountered would be more restricted, simply because less pathophysiological knowledge was activated. The data generally confirmed this hypothesis. The less time available for studying the endocarditis case, the shorter the pathophysiological explanations. In addition, an intermediate effect related to level of expertise was found in the pathophysiology data as well. Students of higher levels have acquired more elaborate knowledge of pathophysiology than students of lower levels or laypersons, hence their pathophysiological explanations of the endocarditis case are more extensive, more accurate, and are at a more detailed level.

The performance of the internists was quite different. Their pathophysiological explanations were short and independent of the amount of time provided for processing the case. Boshuizen and Schmidt (1992) suggest that experts' knowledge of pathophysiology has become encapsulated as a result of practice—hence the short explanation protocols. A direct assessment of the number of encapsulating propositions used in the pathophysiology protocols (Figure 5) confirmed this hypothesis. The amount of encapsulation turned out to be a linear function of experience. Two examples may illustrate the differences in pathophysiological explanations given by experts and advanced students. Figure 6 displays a pathophysiological network constructed on the basis of the explanation protocol of a fourth-year medical student in the 3.5-min processing condition. The protocol consisted of 42

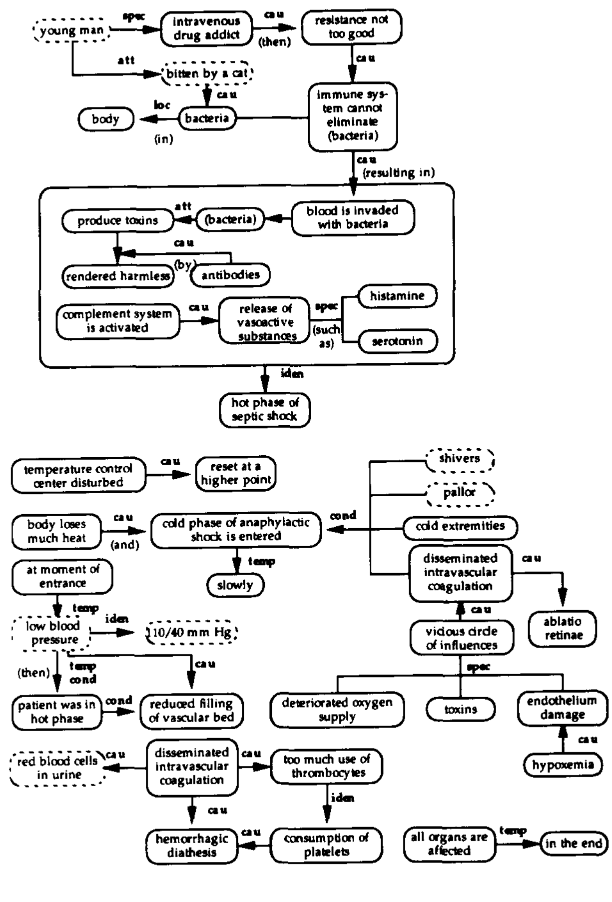


Figure 6. Causal network of propositions derived from Intermediate 87-255's pathophysiological protocol. Dashed frames indicate concepts in the pathophysiology protocols that were directly derived from the endocarditis case. Solid frames indicate concepts that were not. Note that each proposition in the network consists of two concepts or nodes connected by a link or qualifier. These links could be causal (cau), conditional (cond), temporal (temp), attributional (att), or locational (loc). In addition, they could indicate that the second node is a specification of the first (spec). Other qualifiers are negation (neg), identity (iden), and class relation (isa).

propositions. Propositions can share arguments with other propositions, so a protocol can be rewritten as an associative network. For instance, the first propositions of the protocol of Intermediate 87-255 were: "Probably the young man is an intravenous drug addict. He has also been bitten by a cat. His resistance is not too good, so his immune system has not been able to sufficiently eliminate the bacteria which the cat has introduced into the man's body by his bite." Figure 6 shows how these propositions were transformed into an associative network.

Inspection of Figure 6 reveals how the fourth-year student tries to account for almost every sign and symptom embedded in the text of the endocarditis case, suggesting that extensive processing has taken place. Explanation is at a detailed level; quite a few of the propositions include basic physiological mechanisms.

Figure 7 displays the network representation of a randomly selected protocol of an internist with 11 years of experience. The contrast with the intermediate's network is striking—10 propositions were produced. These propositions include only a few cues provided by the case, and the explanation is highly encapsulated. There is no reference to causative agents such as bacteria, or to mediating processes such as sepsis. These seem to be encapsulated in the few highly condensed concepts provided (e.g., endocarditis of the aortic valve) and their relationships. The explanations provided by the other experts have essentially the same structure.

So far, we have only been able to demonstrate that changes in recall of a clinical case under various levels of expertise and time constraints are associated with similar patterns in the post hoc pathophysiological explanations. The hypothesis as proposed, however, is a *causal* one: Subjects of higher levels of expertise, especially when they have sufficient time available, activate more knowledge of pathophysiology in the process of comprehending the case. The extensiveness of their processing determines the elaborateness of the case representation, which in turn determines how much is recalled. This explanation applies to novices and intermediates, but not to experts, because their performance is less dependent on extensive, detailed pathophysiological processing of the case. However, if the processing of pathophysiological knowledge is, indeed, causally responsible for the amount of recall by novices and intermediates, it must be possible to experimentally generate an intermediate effect in free recall by actively manipulating the extent of pathophysiological processing. This was attempted in Experiment 2.

EXPERIMENT 2

In Experiment 2, the priming of prior knowledge paradigm developed by Bransford and colleagues (e.g., Bransford & Johnson, 1972; Franks, Bransford, & Auble, 1982) was used. In this experimental paradigm, prior knowledge is either fully or partially activated; subsequently, subjects engage in a relevant criterial task. Differences in performance, that is, differences in recall of

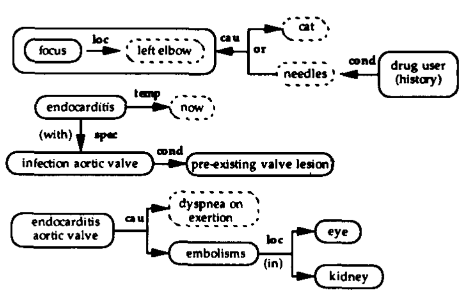


Figure 7. Causal network of propositions derived from Expert 8-023's pathophysiology protocol.

a text, can then be attributed to the nature and amount of prior knowledge activated. In the present experiment, subjects' knowledge of pathophysiology was primed by having them recall whatever they knew about endocarditis (the disease underlying the case used in Experiment 1) for either 30 sec or 3.5 min. Subsequently, both groups studied the endocarditis case for only 30 sec. If case recall depends on the amount of prior knowledge activated, then differences in the activation time allotted will result in recall differences. It was expected that students in the 3.5-min condition would recall more information from the case than students in the 30-sec condition, although case-processing time was constant for all the groups. Such a finding would demonstrate that, for these levels of expertise, recall performance is indeed causally related to the amount of activation of elaborate pathophysiological knowledge, as was suggested by Experiment 1. However, because physicians are less dependent on extensive activation of pathophysiological knowledge to comprehend and recall a case (simply because they operate upon structures that contain pathophysiological knowledge in encapsulated mode), one would expect them to recall less than students. Consequently, an inverted U-shaped relation between expertise level and recall is expected; this time, however, generated experimentally. In addition, the experts should be less affected by the experimental manipulation. Such findings would further support the notion that the intermediate effect in clinical case recall is indeed caused by the activation of elaborate, causal pathophysiological knowledge by students of intermediate levels of expertise. In addition, it would demonstrate that experienced physicians process clinical information using knowledge structures distinctively different from those of intermediates.

Method

Subjects. The subjects were 20 first-year health sciences students and 60 medical students—20 second-, 20 fourth-, and 20 sixth-year undergraduate students. In addition, 20 internists participated in the experiment.

Materials. The endocarditis case that was used in Experiment 1 was used here as well.

Procedure. Each group was randomly assigned to one of two experimental conditions. The subjects were tested individually. Depending on the condition, they were given the opportunity to activate their knowledge of endocarditis for either 3.5 min or 30 sec. To check whether activation of prior knowledge took place in the way intended, the sessions were audiotaped. After completing the activation task, all the subjects were requested to study the acute bacterial endocarditis case for 30 sec. (Prior to the experiment, the subjects were given the opportunity to read an unrelated text of exactly the same length to provide them with experience in scanning a text in a very short time. This was done to minimize variation in the way they would undertake the criterion task.) Subsequently, they were asked to write down whatever information they recalled from the case and to state a most likely diagnosis. The subjects were free to use as much time as they needed for these assignments. The verbatim transcripts of the activation task and the free-recall protocols were segmented into propositions in the same way as described for Experiment 1. The number of propositions produced during priming of prior knowledge was counted. In addition, the number of propositions correctly recalled from the case was established.

Results and Discussion

Diagnostic accuracy. Diagnostic performance was again strongly related to expertise [$F(4,90) = 22.878$, $MS_e = .69$, $p < .0001$]. The amount of time available for activation of prior knowledge had no influence on performance. Figure 8 shows the average diagnostic accuracy scores for each level of expertise and for the two treatment conditions.

Activation of pathophysiological knowledge. Figure 9 displays the number of propositions produced by the students and the physicians in response to the request that they tell everything they knew about endocarditis. As was expected, the subjects were able to produce more propositions when entitled to talk about the subject for 3.5 min as compared with only 30 sec [$F(1,90) = 168.506$, $MS_e = 87.75$, $p < .0001$].

These data illustrate that the experimental manipulation was successful: The more time available, the more prior knowledge was activated. In addition, an effect of exper-

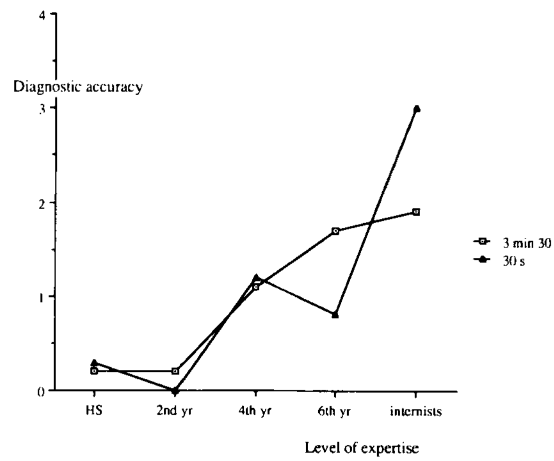


Figure 8. Average accuracy of diagnoses as a function of expertise and level of priming.

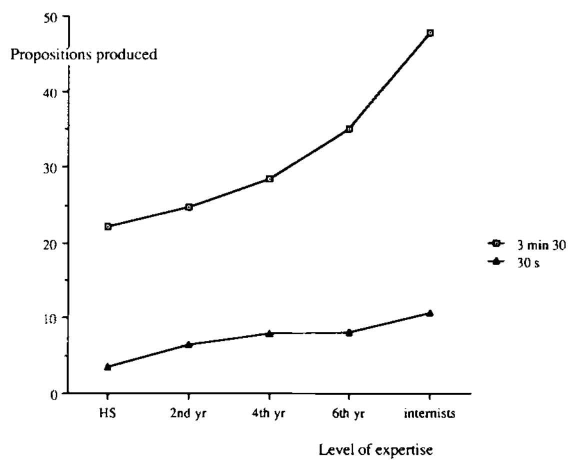


Figure 9. Average number of propositions produced in the priming task.

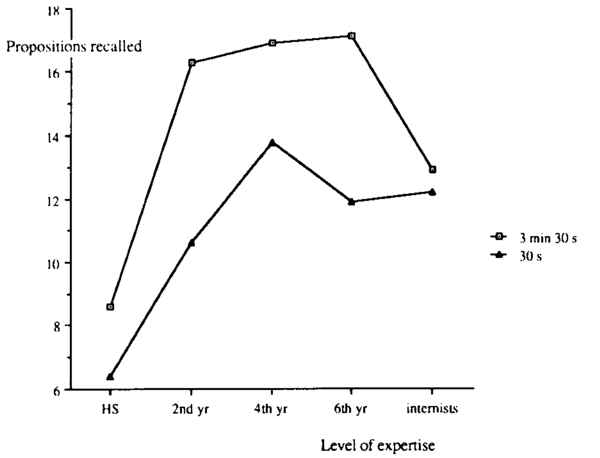


Figure 10. Average number of propositions recalled from the endocarditis case under two levels of priming.

tise is demonstrated [$F(3,90) = 9.168$, $MS_e = 87.75$, $p < .0001$]; the internists generally produced more propositions in response to the priming task than the advanced students. The latter, in turn, generated more knowledge than the novices and laypersons.

Free recall. The prediction was that, among novices and advanced students, the amount of relevant knowledge activated by the priming procedure would influence the comprehension of the clinical case and, hence, its recall. Figure 10 displays the average number of propositions recalled from the endocarditis case when it was subsequently presented to all the subjects for 30 sec.

The free-recall data demonstrate both an effect of prior activation [$F(1,90) = 13.37$, $MS_e = 21.36$, $p < .001$] and an effect of expertise [$F(4,90) = 8.87$, $MS_e = 21.36$, $p < .0001$]. Although all the subjects had only 30 sec to process the case, those who had less opportunity to activate relevant knowledge in advance recalled significantly less from the endocarditis case. Exceptions were the internists [$F(1,18) = .11$, $MS_e = 21.92$, $p < .75$] and the health sciences students [$F(1,18) = 1.77$, $MS_e = 13.71$, $p < .21$]. More important, the data display an intermediate effect, because the effect of expertise on case recall shows a significant quadratic component [$F(1,95) = 20.3533$, $MS_e = 24.15$, $p < .0001$] without significant deviations. Thus, the more advanced students recalled more than both the novices and the experts. As predicted, the priming of prior pathophysiological knowledge had its largest impact on case processing by the students of intermediate levels of expertise, whereas the experts' performances were not affected. These data provide unequivocal evidence for the notion that intermediates' comprehension of a clinical case is driven by detailed pathophysiological knowledge, while experts' understanding is independent of elaborate processing of this kind.

GENERAL DISCUSSION

According to current theories of text processing (Graesser & Clark, 1985; Kintsch, 1988; Voss & Bisanz, 1985), processing of a clinical case by subjects of different levels of expertise, and under various time constraints, would produce two free-recall phenomena. The first would be an increasing recall as a function of expertise; the second would be a decrease in recall when time to study the case becomes more restricted. The general idea here is that reading the case induces the reader to build a cognitive model of the situation to which the text is referring (Van Dijk & Kintsch, 1983). The more knowledge relevant for understanding the case the reader has, and the more time for processing, the more comprehensive the mental representation of the text will be. Since free recall reflects this representation, one may predict that it would increase with expertise and processing time. This prediction agrees with findings in the general text-processing literature (e.g., Spilich et al., 1979) as well as with results of studies in the field of expertise and its development (e.g., Chase & Simon, 1973; de Groot, 1946).

The findings presented in this article, as well as those of others (e.g., Hassebrock et al., 1988; Muzzin et al., 1983; Patel & Groen, 1991), are at variance with these predictions. In Experiment 1, the intermediate groups generally recalled the case better than both experts and novices. These results suggest that predictions derived from theories of text comprehension can account for novice and intermediate performance, but not for expert recall of a clinical case.

In the introduction to this article, we proposed several mechanisms that may explain these results. In line with notions put forward by Kintsch (1988), it was assumed that novices and intermediates, because they have no, or only limited, knowledge fit to the task at hand, would process a clinical case in a bottom-up fashion, extensively activating prior knowledge to arrive at a more or less coherent representation of the case. It was conjectured that such construction-integration processes would involve extensive inferencing and elaboration and would take considerable time. Experts, on the other hand, may have seen many patients with signs and symptoms similar to the one presented and may have developed an appropriate schema or situation model (Van Dijk & Kintsch, 1983; Groen & Patel, 1988) for the task at hand. We have argued that the notion of processing guided by a situation model is in itself insufficient to explain why the experts recalled less information from the case. What must be explained is which structural characteristics enable a situation model to condense the information from a text. Therefore, the idea of knowledge encapsulation was introduced. Encapsulation was defined as the progressive subsumption, or packaging, of lower level concepts and their relations in an associative net under a limited number of high-level concepts with the same explanatory power, which results from repeated application of knowledge to similar situations (Boshuizen & Schmidt, 1992; Schmidt & Boshuizen, 1992, in press). Knowledge in encapsulated mode would cause sets of individual items in a text to be perceived as integrated wholes and would lead to processing this information in an encapsulated format. This would explain why experts' recall is sparse, contains many knowledge-based inferences, and is independent of time constraints on processing. Both recall and pathophysiology data supported the hypothesis that experts use knowledge in encapsulated mode. Experts recalled the case mainly using concepts not found in the text itself, but subsuming most of the information from that text. In addition, their pathophysiological explanations, although also sparse, contained more encapsulating concepts than those of the students.

Experiment 2, in addition, demonstrated that, among students, the free recall of a clinical case is indeed causally determined by the amount of activation of knowledge of pathophysiology, whereas among clinicians it is not. Based on the notions outlined in this article, we were able to experimentally generate an intermediate effect in the recall of a case by manipulating the amount of activation of underlying pathophysiological knowledge. Whereas the students were greatly affected by the conditions of the experiment, the internists processed the patient information largely unaffected by the priming event. Even when the experts were forced to activate knowledge at a detailed level, this knowledge did not affect their comprehension of the clinical case. These findings indicate that the internists superior diagnostic performance cannot be explained by more extensive pathophysiological processing of a case, as is suggested elsewhere (e.g., Lesgold, 1984), but is based on knowledge structures qualitatively different from those of advanced students.

Several objections may be raised against the findings presented here. The first is that intermediate effects may have occurred not because the intermediates and the experts employed qualitatively different knowledge while processing the case, but because they employed different *strategies* (Norman, Brooks, & Allen, 1989). For instance, if the intermediates attempted to simply learn the material presented by heart, whereas the experts processed the case in a meaningful way, this may result in more elaborate free recall by the intermediates, particularly when they had sufficient time to do so. There are several reasons why such differences in processing strategy are unlikely. First, the subjects were not aware beforehand that they would be requested to recall the case. Experiment 1 was presented to them as a diagnostic task, so recall can thus be considered a by-product of an attempt to understand the cause of the signs and symptoms presented in the text. Second, a learning-by-heart explanation would render the post hoc pathophysiological data meaningless. If the students did not activate pathophysiological explanations in the course of processing the case, but just learned the text in a rote-memorization fashion, then where do these differences in pathophysiological explanations related to expertise and processing time come from? And third, a rote-memorization hypothesis certainly cannot account for the results of Experiment 2, because all the groups in this experiment had only 30 sec to study the endocarditis case.

A second objection may be that the data represent a case of output editing, that is, the experts were less motivated to comply with the requirements of the experiments than the students and, therefore, produced shorter protocols. However, this is unlikely, because in Experiment 2, the experts, when asked to state everything they knew about endocarditis, produced, in fact, longer protocols. There is, therefore, no particular reason to believe that the experts were less motivated than the other subjects. A further possibility is that experts are trained in a kind of shorthand to represent cases, whereas students are not. This is a more serious option, because physicians regularly report about the patients they see, whereas students do not have the same experience to the same extent. However, intermediate effects have been reported under conditions in which no post hoc recall was required, and hence no possibility for output editing (Grant & Marsden, 1988; Patel et al., 1988; Patel & Groen, 1991). In addition, physicians' protocols are not only shorter, but also *different* in terms of their content. There is no particular reason to assume that differences in the concepts used for recalling and explaining cases do not reflect differences in the structure underlying performance.

A final objection that may be raised is that free recall may not reflect the problem representation in the medical domain to the same extent that it does in other domains such as chess or computer programming, because its results are so different from the results from these other domains (e.g., Chase & Simon, 1973; Shneiderman, 1976). This issue has troubled many investigators confronted with data similar to ours (e.g., Norman et al., 1989). The present experiments suggest, however, that free recall *is* a meaningful indicator of the problem representation resulting from attempts to comprehend a case. The present studies have demonstrated that it is possible to replicate the findings of de Groot (1946), Chase and Simon (1973), and others in the medical domain, provided that processing time is sufficiently short. The results acquired under the 30-sec processing condition in Experiment 1 illustrate this position. However, the relationship between prior knowledge of a domain and problem representation in that domain may not be as straightforward as has been assumed in previous studies. If the development of expertise involves not only quantitative growth of knowledge, but also qualitative shifts, as is suggested by some (e.g., Feltovich, Johnson, Moller, & Swanson, 1984) and demonstrated in the present experiments, then representation of problems by subjects at different stages of development will tend to reflect these shifts.

A few concluding remarks regarding knowledge encapsulation follow. Encapsulation should not be conceptualized as a permanent condition of the experts' knowledge base, but rather as just one of the possible forms in which knowledge may present itself. This idea fits well with Kintsch's assumption that knowledge organizes itself depending on the task at hand. If the task is understanding or explaining some well-known or routine aspect of the world, knowledge in encapsulated mode will do. Other tasks, for instance, the solving of unfamiliar problems or the requirement to state everything one knows about a topic, may call for the activation of knowledge in its elaborate form. In Experiment 2, it was demonstrated that the experts, when explicitly asked, produced more detailed knowledge about endocarditis than both advanced students and novices. Boshuizen (1989) found similar evidence of "unfolding" of previously encapsulated knowledge. Patel, Groen, and Arocha (1991) provide data that could also be interpreted as evidence for task-dependent knowledge unfolding. They presented medical experts with cases within and outside their immediate domain of expertise and required them to explain these cases in terms of their underlying pathophysiology. The resulting explanation protocols of cases outside the domain were longer, more detailed, and contained more biomedical concepts, suggesting that the experts more extensively activated knowledge when confronted with nonroutine problems.

Encapsulation does not necessarily need to be confined to domains such as medicine. It may well be possible that encapsulation in the knowledge base is responsible for shortcut phenomena in other domains as well, in particular, the chunking of information that is typically observed in grand masters in chess (Chase & Simon, 1973), and "step skipping" by experts in geometry problem-solving (Koedinger & Anderson, 1990).

REFERENCES

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. (1989). *De ontwikkeling van medische expertise: Een cognitief-psychologische benadering* [On the development of medical expertise: A cognitive psychological approach]. Haarlem, The Netherlands: Thesis Publishers.

BOSHUIZEN, H. P. A., & SCHMIDT, H. G. (1992). On the role of biomedical knowledge in clinical reasoning by experts, intermediates and novices. *Cognitive Science*, **16**, 153-184.

BOSHUIZEN, H. P. A., SCHMIDT, H. G., & COUGHLIN, L. D. (1987, April). *On-line representation of a clinical case and the development of expertise*. Paper presented at the meeting of the American Educational Research Association, Washington, DC.

BRANSFORD, J. D., & JOHNSON, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning & Verbal Behavior*, **11**, 717-726.

CHASE, W. G., & SIMON, H. A. (1973). Perception in chess. *Cognitive Psychology*, **4**, 55-81.

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.

CHI, M. T. H., GLASER, R., & FARR, M. J. (1988). *The nature of expertise*. Hillsdale, NJ: Erlbaum.

CLAESSEN, H. F. A., & BOSHUIZEN, H. P. A. (1985). Recall of medical information by students and doctors. *Medical Education*, **19**, 61-67.

COUGHLIN, L. D., & PATEL, V. L. (1986, April). *Text comprehension and expertise in the domain of medicine*. Paper presented at the meeting of the American Educational Research Association, San Francisco, CA.

DE GROOT, A. D. (1946). *Het denken van den schaker* [Thinking processes in chess players]. Den Haag, The Netherlands: Noord Holland.

ELSTEIN, A. S., SHULMAN, L. S., & SPRAFKA, S. A. (1978). *Medical problem-solving*. Cambridge, MA: Harvard University Press.

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* (pp. 275-319). Reading, MA: Addison-Wesley.

FRANKS, J. J., BRANSFORD, J. D., & AUBLE, P. M. (1982). The activation and utilization of knowledge. In C. R. Puff (Ed.), *Handbook of research methods in human memory and cognition* (pp. 395-425). New York: Academic Press.

GRAESSER, A. C., & CLARK, L. F. (1985). *Structures and procedures of implicit knowledge*. Norwood, NJ: Ablex.

GRANT, J., & MARSDEN, P. (1988). Primary knowledge, medical education and consultant expertise. *Medical Education*, **22**, 746-753.

GROEN, G. J., & PATEL, V. L. (1988). The relationship between comprehension and reasoning in medical expertise. In M. T. H. Chi, R. Glaser, & M. Farr (Eds.), *The nature of expertise* (pp. 287-310). Hillsdale, NJ: Erlbaum.

HASSEBROCK, F., BULLEMER, P., & JOHNSON, P. E. (1988, April). *When less is more: Selective memory of problem-solving experts*. Paper presented at the meeting of the American Educational Research Association, New Orleans, LA.

HUNTER, I. M. L. (1962). An exceptional talent for calculative thinking. *British Journal of Psychology*, **53**, 243-258.

KINTSCH, W. (1974). *The representation of meaning in memory*. Hillsdale, NJ: Erlbaum.

KINTSCH, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, **95**, 163-182.

KINTSCH, W., & GREENO, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review*, **92**, 109-129.

KOEDINGER, K. R., & ANDERSON, J. R. (1990). Abstract planning and perceptual chunks: Elements of expertise in geometry. *Cognitive Science*, **14**, 511-550.

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* (pp. 31-60). San Francisco: W. H. Freeman.

LESGOLD, A., RUBENSTEIN, 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* (pp. 311-342). Hillsdale, NJ: Erlbaum.

MCKEITHEN, K. B., REITMAN, J. S., RUETER, H. H., & HIRTLE, S. C. (1981). Knowledge organization and skill differences in computer programmers. *Cognitive Psychology*, **13**, 307-325.

MORRIS, P. E., GRUNEBERG, M. M., SYKES, R. N., & MERRICK, A. (1981). Football knowledge and the acquisition of new results. *British Journal of Psychology*, **72**, 479-483.

MUZZIN, L. J., NORMAN, G. R., FEIGHTNER, J. W., & TUGWELL, P. (1983). Expertise in recall of clinical protocols in two specialty areas. *Proceedings of the 22nd Conference on Research in Medical Education* (pp. 122-128). Washington, DC: American Association of Medical Colleges (AAMC).

MUZZIN, L. J., NORMAN, G. R., JACOBY, L. L., FEIGHTNER, J. W., TUGWELL, P., & GUYATT, G. (1982). Manifestations of expertise in recall of clinical protocols. *Proceedings of the 21st Conference on Research in Medical Education* (pp. 163-168). Washington, DC: American Association of Medical Colleges (AAMC).

NEWELL, A., & ROSENBLOOM, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition*. Hillsdale, NJ: Erlbaum.

NEWELL, A., ROSENBLOOM, P. S., & LAIRD, J. E. (1989). Symbolic architectures for cognition. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 93-132). Cambridge, MA: MIT Press.

NORMAN, G. R., BROOKS, L. R., & ALLEN, S. W. (1989). Recall by expert medical practitioners and novices as a record of processing attention. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 1166-1174.

NORMAN, G. R., JACOBY, L. L., FEIGHTNER, J. W., & CAMPBELL, E. J. M. (1979). Clinical experience and the structure of memory. *Proceedings of the 18th Annual Conference on Research in Medical Education* (pp. 214-217). Washington, DC: American Association of Medical Colleges (AAMC).

PATEL, V. L., EVANS, D. A., & GROEN, G. J. (1989). Biomedical knowledge and clinical reasoning. In D. A. Evans & V. L. Patel (Eds.), *Cognitive science in medicine: Biomedical modeling* (pp. 53-112). Cambridge, MA: MIT Press.

PATEL, V. L., EVANS, D. A., & KAUFMAN, D. R. (1988). A cognitive framework for doctor-patient interaction. In D. A. Evans & V. L. Patel (Eds.), *Cognitive science in medicine: Biomedical modeling* (pp. 257-312). Cambridge, MA: MIT Press.

PATEL, V. L., & GROEN, G. J. (1986a). Knowledge-based solution strategies in medical reasoning. *Cognitive Science*, **10**, 91-116.

PATEL, V. L., & GROEN, G. J. (1986b, November). *Non-monotonicity in novice-intermediate-expert comparisons*. Paper presented at the meeting of the Psychonomic Society, New Orleans, LA.

PATEL, V. L., & GROEN, G. J. (1991). Developmental accounts of the transition from medical student to doctor: Some problems and suggestions. *Medical Education*, **25**, 527-535.

PATEL, V. L., GROEN, G. J., & AROCHA, J. (1991). Medical expertise as a function of task difficulty. *Memory & Cognition*, **18**, 394-406.

PATEL, V. L., & MEDLEY-MARK, V. (1985). Knowledge integration from clinical texts, use of factual, inferential and integrative questions. *Proceedings of the 24th Annual Conference on Research in Medical Education* (pp. 18-22). Washington, DC: American Association of Medical Colleges (AAMC).

REITMAN, J. S. (1976). Skilled perception in Go: Deducing memory structures from inter-response times. *Cognitive Psychology*, **8**, 336-356.

SCHMIDT, H. G., & BOSHUIZEN, H. P. A. (1992). Encapsulation of biomedical knowledge. In D. A. Evans & V. L. Patel (Eds.), *Advanced models of cognition for medical training and practice*. New York: Springer-Verlag.

SCHMIDT, H. G., & BOSHUIZEN, H. P. A. (in press). On acquiring expertise in medicine. *Educational Psychology Review*.

SHNEIDERMAN, B. (1976). Exploratory experiments in programmer behavior. *International Journal of Computer & Information Services*, **5**, 123-143.

SPILICH, G. J., VESONDER, G. T., CHIESI, H. L., & VOSS, J. F. (1979). Text processing of domain-related information for individuals with high and low domain knowledge. *Journal of Verbal Learning & Verbal Behavior*, **18**, 275-290.

VAN DIJK, T. A., & KINTSCH, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.

VOSS, J. F., & BISANZ, G. L. (1985). Knowledge and the processing of narrative and expository texts. In B. K. Britton & J. B. Black (Eds.), *Understanding expository text* (pp. 173-198). Hillsdale, NJ: Erlbaum.

NOTES

1. These consequences of Kintsch's (1988) model for expertise development have to be inferred, because Kintsch does not address the issue of novice-expert differences in text processing.

2. A distinction is made here between clinical and biomedical knowledge. Clinical knowledge is defined as knowledge of the attributes of sick people. It refers to 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, or pathophysiological knowledge, by contrast, refers to the pathological principles, mechanisms, or processes underlying the manifestations of disease. It is phrased in terms of entities such as viruses or bacteria, and in terms of tissue, organs, organ systems, or bodily functions.

3. Knowledge encapsulation has features in common with two other knowledge-"packaging" cognitive processes described in the literature. Anderson's (1983, 1987) ACT theory assumes that, with practice, procedural knowledge, expressed as a set of productions, collapses into a fewer number of more comprehensive productions. This process is called *knowledge compilation*. A similar learning mechanism, called *chunking*, is described by Newell and Rosenbloom (1981). According to Newell, Rosenbloom, and Laird (1989), both compilation and chunking serve the purpose of tuning the knowledge base, that is, generalizing or specializing productions. However, whereas knowledge compilation and chunking result from the reduction of a set of productions in a *procedural* knowledge base into a smaller number of more encompassing ones, encapsulation operates merely upon the declarative knowledge base, gradually reducing, through practice, molecular concepts and their interrelations to a smaller set of molar ones.

4. In contrast to the situation in the USA, Dutch medical students follow a 6-year undergraduate program.

5. In plain English, the following explanation applies to the case: The young man is probably a drug addict who has used a contaminated syringe. Bacteria have caused a sepsis, to which the body responds with fever. The sepsis condition is responsible for the development of an endocarditis, which is a vegetation or abscess on the aortic valve. This causes aortic insufficiency and thromboembolies. The disease is particularly acute in this patient because he has a decreased bodily resistance as a result of his drug use. The loss of vision and the erythrocytes in his urine are the visible symptoms of the thrombo-embolies, while the low blood pressure, the diastolic murmur, the shortness of breath, and the high pulse rate can be attributed to the aortic insufficiency. Most other symptoms are fever-related. The story of the cat bite is probably a cover-up for the puncture wounds in his left arm.

6. These data seem to explain why an intermediate effect was not demonstrated in all recall studies reviewed in the introduction (e.g., Muzzin et al., 1982; Norman et al., 1979). In these particular studies, processing time was limited. As is demonstrated by the present experiment, the emergence of an intermediate effect depends on the amount of time available for processing.

APPENDIX A

Text of the Acute Bacterial Endocarditis Case (Patel & Groen, 1986a, p. 97)

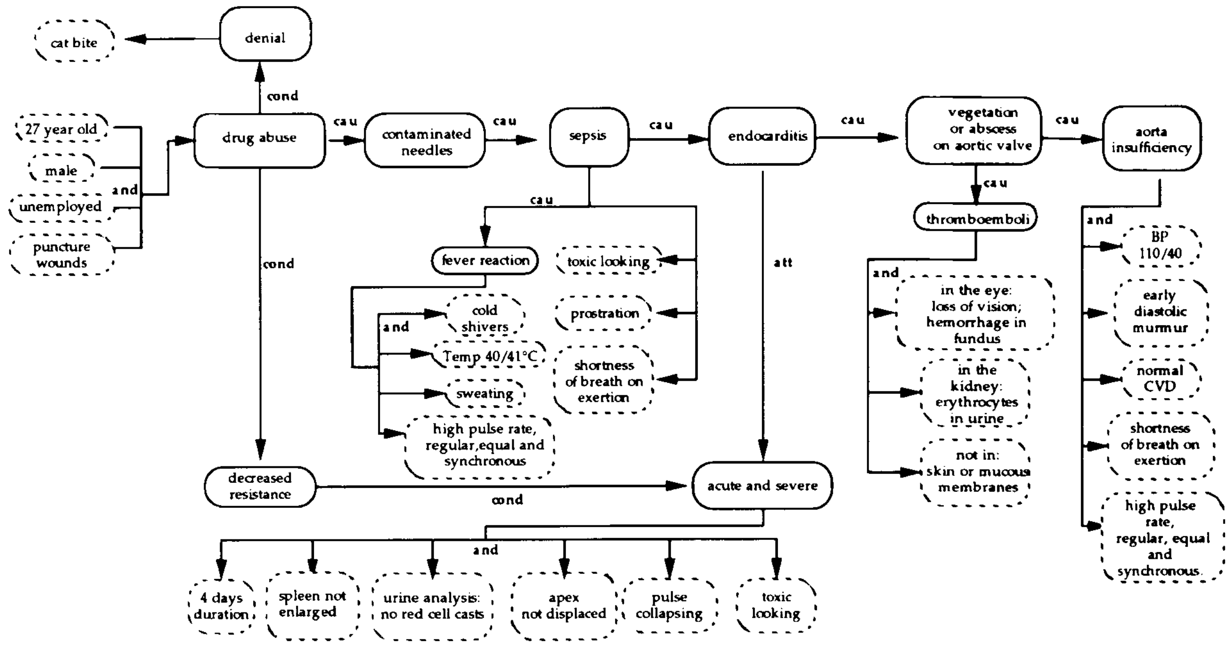
A 27-year-old unemployed male was admitted to the emergency room. He complained of shaking chills and fever of 4 days duration. He took his own temperature and it was recorded at 40° C on the morning of his admission. The fever and chills were accompanied by sweating and a feeling of prostration. He also complained of some shortness of breath when he tried to climb the two flights of stairs in his apartment. The patient volunteered that he had been bitten by a cat at a friend's house a week before admission.

Functional inquiry revealed a transient loss of vision in his right eye, which lasted approximately 45 seconds. This he described the day before admission to the emergency ward.

Physical examination revealed a toxic-looking young man who was having a rigor. His temperature was 41° C. Pulse was 124 per minute. BP 110/40. Mucous membranes were clear. Examination of his limbs showed puncture wounds in his left antecubital fossa. There were no other skin findings.

Examination showed no jugular venous distention. Pulse was regular, equal and synchronous. The pulse was also noted to be collapsing. The apex beat was not displaced. Auscultation of his heart revealed a 2/6 early diastolic murmur in the aortic area. Funduscopy revealed a flame shaped hemorrhage in the left eye. There was no splenomegaly. Urinalysis showed numerous red cells. There were no red cell casts on microscopic urinalysis.

APPENDIX B
Canonical Explanation for the Endocarditis Case



Note—The concepts enclosed by solid frames, and their interrelations, constitute the minimally sufficient canonical explanation. Dashed frames contain the signs and symptoms displayed in the endocarditis case explained by the model. For an explanation of the meaning of the links between concepts, see the caption for Figure 6.

(Manuscript received February 4, 1991;
revision accepted for publication August 21, 1992.)