## Title
Specificity of Expertise in Clinical Reasoning



**Journal**
Proceedings of the Annual Meeting of the Cognitive Science Society, 8(0)

**Authors**
Joseph, Guy-Marie
Patel, Vimla L.

**Publication Date**
1986

Peer reviewed

# Specificity of Expertise in Clinical Reasoning

**Guy-Marie Joseph & Vimla L. Patel**
**McGill University**

An important development in recent research in problem solving has been the development of new theories and more rigorous methods with which to formulate detailed hypotheses about the processes that constitute expert performance in various domains. Domains in which such research can be found include chess (Chase & Simon, 1973), physics (Larkin et al., 1980), and medicine (Elstein et al., 1978; Patel & Groen, 1986).

Studies of expertise have led to the hypothesis that the dominant component of expert performance is related to the expert's ability to represent problems successfully, and that this results from the expert's having a rich and well organized knowledge base in which patterns of features in the problem are associated with concepts at varying levels of generality. This enables experts to recognize the salient features of the problem, to interpret information relevant to its solution (Johnson et al., 1981), and search for the methods and operations to be used in solving the problem (Greeno & Simon, 1984) all with greater efficiency than novices. Moreover, expert problem solvers tend to use forward chaining strategy when solving a simple or familiar problem (Simon & Simon, 1978; Chi et al., 1981; Glaser, 1985; Larkin et al., 1980). This forward-chaining strategy changes, however, to a very sophisticated means- ends analysis (Greeno & Simon, 1985; Chi et al., 1981) or backward chaining (Bhaskar & Simon, 1977) when solving a difficult or unfamiliar problem. These studies have provided evidence suggesting hypotheses about the kinds of processes and knowledge structures experts use while working on familiar problems or problems within their domain of expertise.

These research suggests that expertise is determined by subjects' (i) prior knowledge of a particular domain, (ii) their knowledge of general problem solving methods, and (iii) their knowledge of any domain-specific heuristics. Testing experts in problem solving outside their domain of expertise should thus make it possible to contrast their performance in the two domains and thereby begin to assess the relative contributions of domain-specific knowledge vs. general problem solving ability. The need for making this contrast, and the difficulty of doing so, has been pointed out most recently by Greeno & Simon (1984).

Studies of experts working on problems outside their specialty area have found that experts' problem-solving skills diminish notably in such unfamiliar domains (Kassirer et al., 1982). Miller (1975) found that experts approach non-specialty problems "more parsimoniously", and that they resort to strategies commonly associated with novice performance. Means-ends analysis was found to be the primary strategy used when the problem solver has little domain-specific knowledge (Greeno and Simon,1984). Ernst and Newell (1969) have shown that with unfamiliar problems, experts rely more on problem-solving strategies rather than on domain-specific knowledge. Boshuizen and Claessen (1977) identified some of the problems involved in such studies. Another important characteristics of expert performance is the difference between the solution strategies used by experts and novices. The distinction is that experts use a "forwar-chaining" strategy, whereas novices use a "backward-chaining strategy".

Subjects' problem spaces are determined, in general, by the search methods that the subjects use, the features of the problem's solution, and their knowledge of the problem domain. In situations where expert problem solvers do not have special training in a problem domain, they must construct a problem space that includes representations of the strategies and constraints necessary to solve the problem. If subjects have special training or experience in the problem domain, their prior knowledge includes general characteristics of the problem space, and their representations of the individual problems are based on this knowledge (Greeno and Simon, 1984). These results and observations lead us to believe that experts working within and outside of their specialities will use very different strategies and/or whatever domain knowledge they have available.

Attempts to assess the relative contributions of domain knowledge vs. problem solving strategies require specific information about both how the subject's representation of the problem changes over time and the domain knowledge necessary to construct and modify these representations. Neither of the two most widely used methods, however, seem to be able to provide this kind of information (Patel et al., 1986). "Think-aloud" protocols (e.g. Greeno & Simon, 1984; Kassirer et al., 1982) provide rich, complex data that are approximately concurrent with the subject's reasoning and therefore provide information about the subject's changing representation of the problem. However, protocol analysis methods (e.g., Ericsson & Simon, 1984) have been limited in their success at providing more than global information about subjects' processing. One of the major concerns raised by protocol analysis is the lack of formal, objective methods for analyzing complex verbal material. Methods of propositional analysis have been proposed to meet this concern (e.g., Kintsch, 1974; Frederiksen, 1975) and have most often been applied to stimulus texts and recall protocols. Although recall protocols provide quite detailed data about representation in memory and inferences based on domain knowledge, they are very limited in the amount of information they can provide about how representations of a problem change over time.

Recently, a number of researchers have addressed these problems by combining the two methods, i.e., performing propositional analyses of think-aloud or on-line protocols (e.g. Kintsch & Greeno, 1985). This strategy has proven useful for studying strategies of reading comprehension (Scardamalia & Bereiter, 1984) and writing (Breuleux & Bracewell, 1986). As a further refinement, this study prompts subjects to think aloud at regular intervals. This on-line procedure allows finer control of the stimuli the subject is to respond to, and may provide more detailed information about the subject's changing representation of the problem as a function of specific portions of the input text.

In the present study we are primarily interested in testing the extent to which expertise is dependent on domain knowledge, rather than the interaction of domain knowledge and problem solving processes. If expertise is primarily dependent on domain knowledge, then when working on problems outside their domain of expertise, experts should resort to strategies (e.g. backward chaining) that are commonly associated with novices. However, if expertise is more a function of general problem solving methods, then there should be few differences in expert performance across the boundaries of their domain of expertise.

## METHOD

*Subjects.*

Nine senior physicians associated with the Faculty of Medicine at McGill University volunteered as subjects for the study. Although all subjects had a general knowledge of medicine, the two groups differed in their specific knowledge of endocrinology: there was a high domain-knowledge (HDK) group consisting of four endocrinologists and a low domain-knowledge (LDK) group consisting of five cardiologists.

*Materials.*

A description of a clinical case in endocrinology based on a real patient was selected as stimulus material (Appendix 1). The case was described in the following order; the patient's history (Segments 1 to 4), findings obtained from examination of bodily systems (Segments 4 to 22), and results from laboratory tests (Segments 23 and 24).

The diagnosis of the case has three subcomponents varying in specificity. The first subcomponent is a slow progressive decrease in thyroid function: chronic hypothyroidism. The second subcomponent is a change from a chronic to an acute form of hypothyroidism, precipitated by a prescribed cough medicine and leading to a state of myxedema pre-coma. The third subcomponent is an autoimmune condition called Hashimoto's thyroiditis as the cause of the hypothyroidism. The complete diagnosis, then, is: Hashimoto's thyroiditis leading to hypothyroidism, precipitated to myxedema pre-coma by previous medication.

*Procedure.*

Subjects were tested individually at their hospital offices. They were first given a short practice session to familiarize themselves with the procedure. The clinical case was presented on the screen of an Apple Macintosh one sentence at a time, using software developed for this purpose by the first author. Subjects initiated the presentation of each sentence by pressing the mouse button. For each sentence presented, subjects were asked to verbalize their thoughts about the contribution of the sentence towards the generation of a diagnosis. As soon as the subject pressed the mouse button again, the next sentence was presented, replacing the previous one. Subjects were not allowed to look back to any prior part of the text. After presentation of the entire case, subjects were asked to summarize the case and to provide their diagnosis(es). Subjects' verbalizations were recorded and later transcribed verbatim.

## ANALYSES

*Protocol Analysis.*

The information in each text segment of the original stimulus text was initially categorized as either critical, relevant, or nonrelevant, based on the judgements of two expert endocrinologists. Cues were rated as "critical" if they were deemed necessary and sufficient for the generation of the accurate diagnosis. Relevant cues were considered necessary but not sufficient for the generation of the accurate diagnosis. Nonrelevant cues included the remainder of the information, the omission of which should not have affected the generation of an accurate diagnosis. The inter-rater reliability was .92 for the critical cues and .85 for the relevant cues.

Each subject's protocol was transcribed so that each segment from the original text was followed by the physician's comments on that segment. Subjects' comments on each segment were then coded for the presence of the different types of cues, and links generated from the cues to hypothetical causes. If the cue was only repeated and no link was suggested, then it was coded as absent (see example 1). If the cue was repeated and a link was suggested, then the cue was coded as accessed and a link was coded as hypothesized (see example 2).

(1) ***Segment #9***: Her body temperature was 36 deg. C.
***Subject***: Her body temperature was 36 deg. C., that does not tell me very much.

In the preceding segment (see Appendix 2, Subject 2), there are no cues coded as accessed and no links generated from it.

(2) ***Segment #3***: She complained of feeling tired all the time, had a loss appetite, a 30 lb. weight gain and constipation.

***Subject***: She had a loss of appetite, a 30 lb. weight gain, and constipation. OK... Right now we are wondering whether *she has got hypothyroidism* when you are looking at this.

In this situation there are three cues accessed (underlined), and three links generated from these cues (linking statement in *italics* ) to one of the subcomponents of the diagnosis (hypothyroidism). This method is described in more detail in Joseph & Patel (1986).

*Accuracy of diagnoses.*

The diagnoses of all the subjects were analyzed for the presence or absence of the three subcomponents with the assistance of an expert endocrinologist who did not know which subjects belonged to which group.

*Frame analyses.*

One representative subject from each group was analysed in more detail (Subjects 2 and 3). Frame analyses (cf. Patel & Groen, 1986) based on methods of propositional analysis were performed on the relations that the subjects generated between cues and their hypothetical causes, and generated semantic networks corresponding to each hypothesized relation. Subjects' frames were then compared to determine their strategies for using the cues provided to generate diagnostic hypotheses, as well as differences in the problem representations they generated as the information was presented.

## RESULTS AND DISCUSSION

*Protocol analysis: Cues selected and links generated.*

The results of the analysis of cues selected and links generated indicate that the main difference between the two groups is the number of links generated from the critical cues (Figure 1). High domain knowledge HDK subjects generated more links from the critical cues selected whereas low domain knowledge LDK subjects generated fewer links from the critical cues selected. Both groups generated fewer links from relevant cues. However, the HDK group selected more relevant cues and generated more links from these cues than did the LDK group.

There was also greater variability between groups in the relationship between critical cues selected and links generated, than for the links generated from relevant cues.

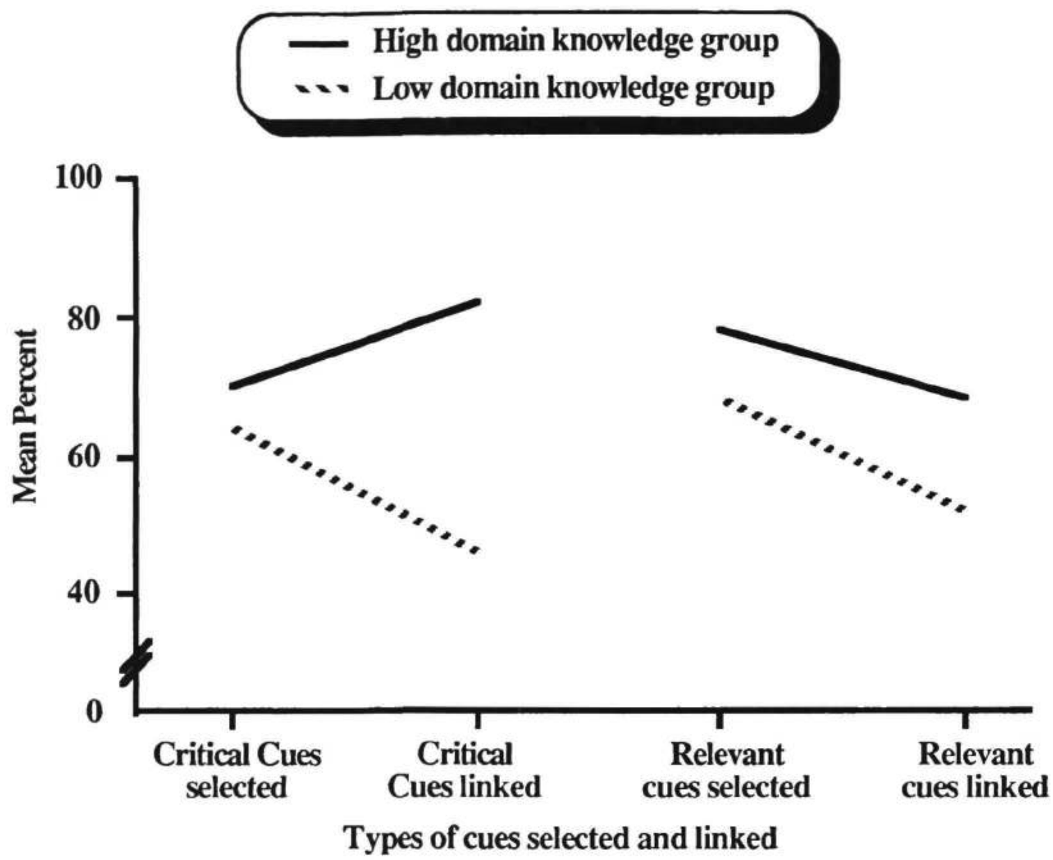


Figure 1. Mean percent of cues selected and linked by high and low domain knowledge subjects.

*Accuracy of diagnoses*:

All the subjects in this group generated the general subcomponent of hypothyrodism in their diagnostic statement (Figure 2). Sixty percent generated the myxedema subcomponent, but they all omitted the pre-comatose aspect of the myxedema. Finally, only twenty percent of the subjects generated the Hashimoto subcomponent as part of their diagnostic statement. In addition, subjects generated differential diagnoses, an ordered list of the most likely diagnostic possibilities of the causes of the patient's disorder. The most frequent differentials listed were: Adrenal failure, pituitary tumor, Hypophyseal disorder, and pericardial effusions.

Thus far, the use of surface measures to evaluate the performances of experts working on problems within and outside their domain of expertise has not yielded meaningful information about the role of domain knowledge and problem solving strategies in the process of clinical reasoning. The main difference in the analysis of accuracy of the diagnosis that LDK experts are more general in determining their diagnostic statement. Their uncertainty thus explains why they generate a very general list of differentials for the causes of the patient's disorder.

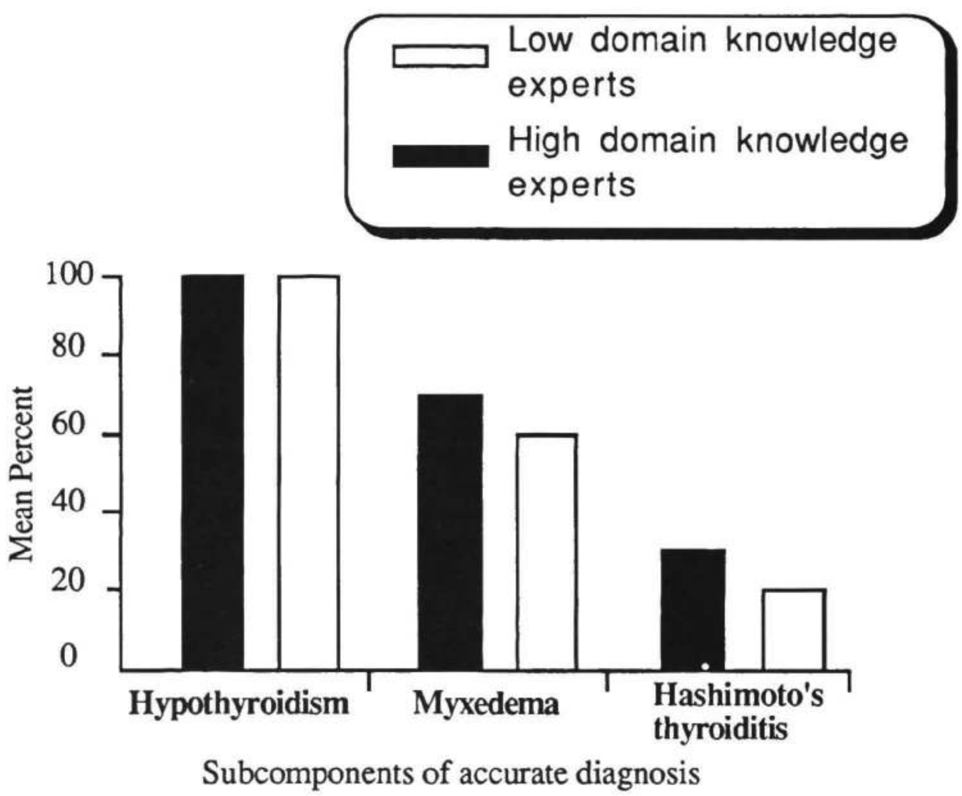


Figure 2. Mean percent of subcomponents of the acurate diagnosis by high and low domain knowledge subjects.

The latter finding is consistent with studies of comprehension of clinical cases in which no differences were found for experts' recall of relevant and non-relevant information in "atypical" cases (Coughlin & Patel, 1986). Therefore, more detailed and structural analysis of, the heuristic strategies and prior knowledge that are needed to generate a representation of the problem is expected to yield process differences between the two groups. The results of this analysis is provided in the next section.

*Analysis of the online protocol of a HDK subject*: The analysis of the structural representations of the online protocol of an HDK subject is presented segment by segment in Figure 3. Segments in which there were no responses were omitted. The first important finding of this analysis was that although this subject identified the three subcomponents of the diagnosis, he did not report them when asked to provide his diagnosis. His diagnostic hypothesis started with a "forward chaining" strategy, followed by a "backward chaining" strategy used to elaborate on the underlying causal mechanisms of the relationship between the cues and the hypothesis (Figure 3, Segments 3 and 4). After the generation of the three diagnostic subcomponents (Segment 7), the processing of the cues followed a model of "pattern recognition" in order to either confirm the diagnostic hypothesis or to rule out the two hypotheses that were generated at the beginning of the case.

*Analysis of the online protocol of a LDK subject*: The analysis of the structural representation of the online protocol of a LDK subject is provided in Figure 4. Contrary to the representations of the HDK subject, this subject provided only the general subcomponent of the diagnosis, hypothyrodism at Segment 4. The other two components were not mentioned with the exception of Segment 9, where the subject mentions the "advanced hypothyrodism". The subject generated many diagnostic hypotheses (e.g. psychiatric disorder, chronic renal disease, anemia) without ruling them out as possibilities, and seemed to generate the hypotheses using a form of "backward chaining" (Segments 4 and 5). Finally, this subject generated very few causal explanations (e.g. causal links) and instead used a process that was more cue driven. That is, evaluated individual cues as they were presented. In some instances he ruled out a possible hypothesis but reinstated the hypothesis at the presentation of the next cue (Segments 10 and 11).

The results of the following study would tend to support the hypothesis that when working on problems in which experts lack specific domain knowledge, they rely on general knowledge in

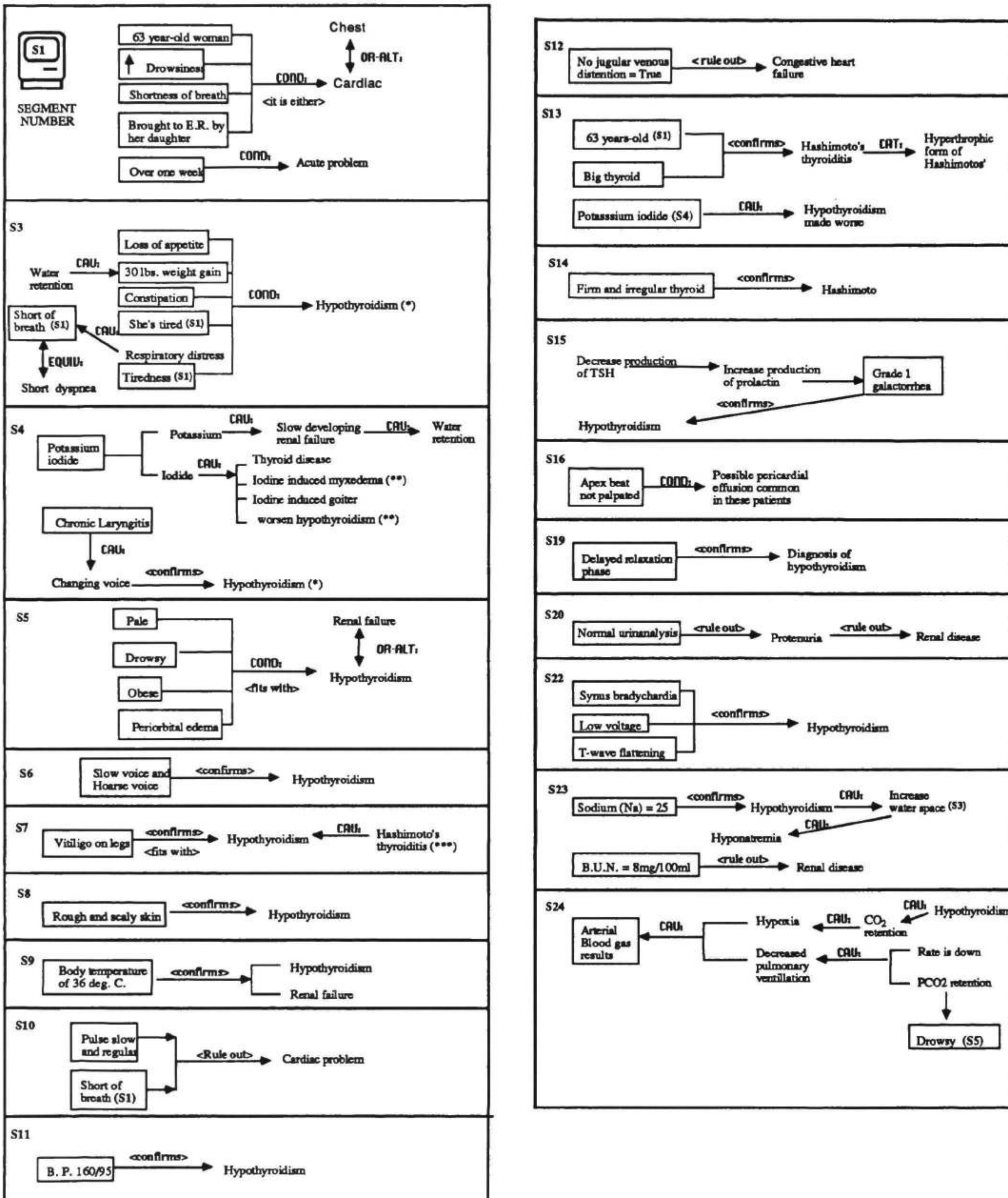


Figure 3. Structural representation of the online protocol segment by segment of a high domain knowledge subject. (S#2)

(*): Indication of the generation of the first subcomponent of the diagnosis for the first time
(**) Indication of the generation of the second subcomponent of the diagnosis for the first time
(***) Indication of the generation of the third subcomponent of the diagnosis for the first time

Text cues

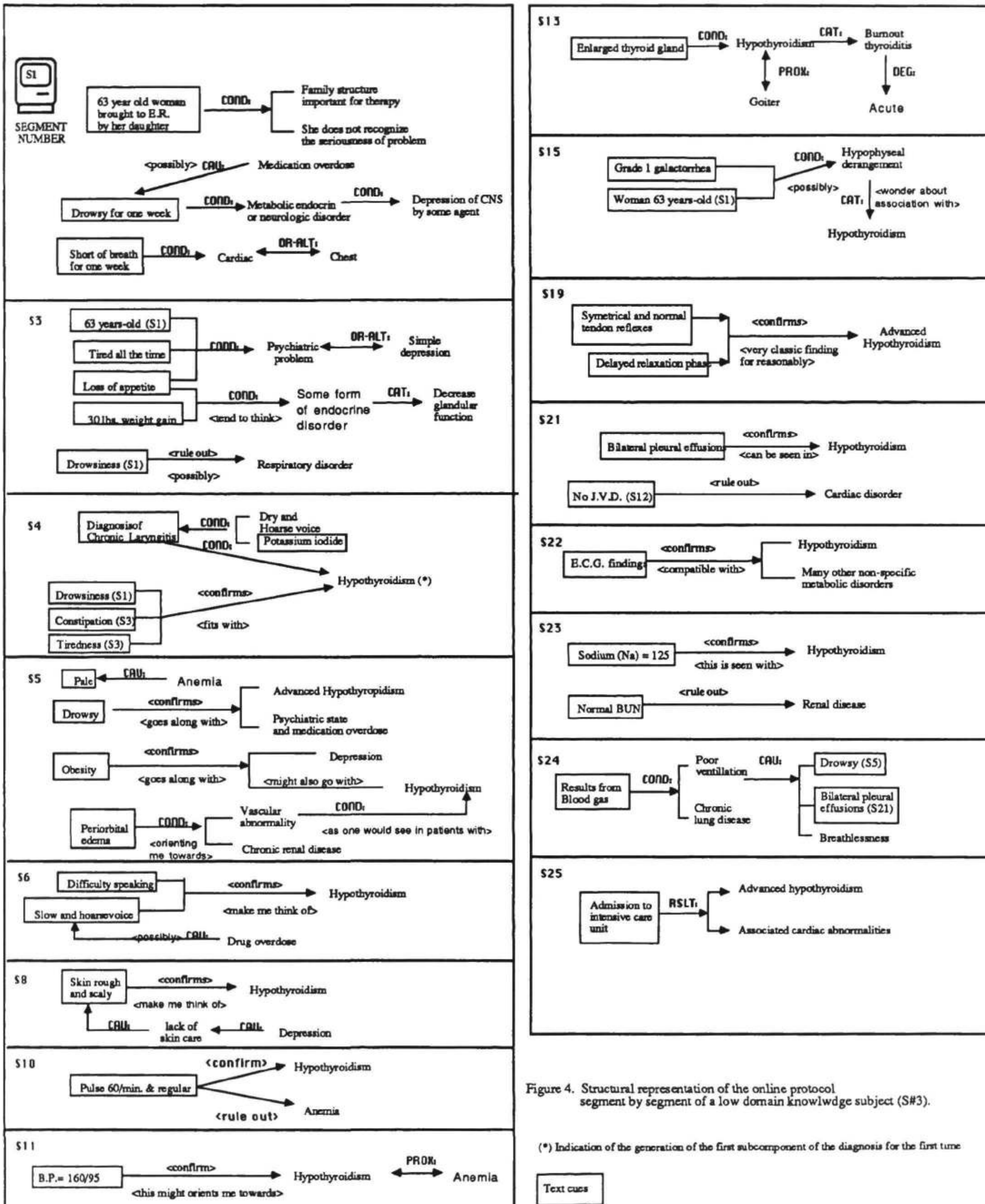


Figure 4. Structural representation of the online protocol segment by segment of a low domain knowlwdge subject (S#3).

(*) Indication of the generation of the first subcomponent of the diagnosis for the first time

Text cues

the problem domain to solve the problem. However, the use of general knowledge directly affects the overall representation of the problem as well as the types of strategies that are used to generate the solution.

The analysis of the two representations and strategies generated by the two subjects revealed information about process differences between HDK and LDK experts. The HDK subject used forward chaining to generate hypotheses and used backward chaining strategy to explain the relationship between the cues and the hypotheses, in providing the underlying pathophysiological mechanisms of the patient's disease. This finding is consistent with the earlier finding that HDK experts generated more links from critical cues than the LDK expert. The HDK subject's representation of the problem was also more elaborated with more causal links. He also provided fewer diagnostic hypotheses and ruled out more alternative hypotheses. These finding led to the use of the term "mechanismic" to describe the process of the HDK expert.

The LDK expert used more backward chaining strategies. He generated more diagnostic hypotheses and his problem representation was much less elaborated and detailed than the HDK expert. The hypotheses generated by the LDK expert were very general and were related to the general subcomponent of hypothyroidism. He evaluated and classified the cues in terms of their relationship to the condition of hypothyroidism. We therefore use the term "classificatory" to describe the process of the LDK expert.

The methods and analysis techniques used in this study reflect the complex nature of the relationship between domain knowledge and problem solving strategies, as well as the problem that researchers must face in the assessment of such complex phenomena. The use of the online methods seems very promising in the study of cognitive processes, especially to study the types of heuristic strategies that used to solve problems, strategiy shifts, and the types of knowledge required to do so.

## ACKNOWLEDGEMENTS

We wish to thank Drs. Y. Patel, S. Magder and D. St. Vil for their assistance with the medical aspects of the study. We particularly want to acknowledge the support of the physicians who generously volunteered their time. We wish to thank M. Dillinger, Dr. R Bracewell, and A. Breuleux, and the graduate students for their help at the various stages of the study. This work was supported in part by grants from the Josiah Macy Jr. Foundation (No. B8520002) and the Natural Sciences and Engineering Council of Canada (A2598) to Vimla Patel.

## REFERENCES


Bereiter, C., & Bird, M. (1985). Use of thinking aloud in identification and teaching of reading comprehension strategies. Cognition & Instruction, 2, 131-156.

Bhaskar, R., & Simon, H. A. (1977). Problem solving in semantically rich domains: An example from engineering thermodynamics. Cognitive Science, 1, 193-215.

Boshuizen, H. P. A., & Classen, H. F. A. (1982). Problems of research into medical-problem solving. Medical Education, 16, 81-87.

Breuleux, A. & Bracewell, R. J. (1986). A discourse analysis approach to the analysis of think aloud protocols. McGill working papers in cognitive sciences #7. Montreal, McGill University.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. Cognitive Psychology, 1, 55-81.

Chi, M. T. H., Glaser, R., & Rees, E. (1981). Expertise in problem solving. In R. J. Sternberg Advances in the psychology of human intelligence. LEA.

Coughlin, L. D., & Patel, V. L. (1986). Text comprehension and the effect of expertise in the domain of medicine. Paper presented at the Annual Conference of the American Education Research Association, San Francisco, CA.

Elstein, A. S., Shulman, L. S., & Sprafka, S. A. (1978). Medical problem solving: An analysis of clinical reasoning. Cambridge, MA: Harvard University Press.

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

Ernst, G. W., & Newell, A. (1969). GPS: A case study in generality and problem solving. New York: Academic Press.

Frederiksen, C. H. (1975). Representing logical and semanticstructure of knowledge acquired from discourse. Cognitive Psychology, 7, 371-458.

Glaser, R. (1985). Thoughts on expertise. Technical report #8, LRDC, University of Pittsburgh, Pittsburgh, PA.

Greeno, J. G., & Simon, H. A. (1986). Problem solving and reasoning. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), Stevens' handbook of experimental psychology (rev. ed.). New York: Wiley.

Groen, G. J., & Patel, V. L. (1985). Medical problem solving and cognitive psychology: Some questionable assumptions. Medical Education 19, 95-100.

Hayes-Roth, B., & Hayes-Roth, F. (1979). A cognitive model of planning. Cognitive Science, 3, 275-310.

Johnson, P. E., Duran, A. S., Hassebrock, F., Moller, J., Prietula, M.,Feltovich, P. J., & Swanson, D. B. (1981). Expertise and error in diagnostic reasoning. Cognitive Science, 5, 235-283.

Joseph, G-M., & Patel, V. L. (1986). Online analysis of expertise in medical problem solving. Technical Report No. 11, Center for Medical Education, Faculty of medicine, McGill University.

Kassirer, J. P., Kuipers, B. J., & Gorry, G. A. (1982). Toward a theory of clinical expertise. The American Journal of Medicine 73, 251-259.

Kintsch, W. (1974). The representation of meaning in memory. Hillsdale, NJ: LEA.

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

Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. Science, 208, 1335-1342.

Miller, P. B. (1975). Strategy choice in medical diagnosis, Project MAC. Report no. TR-153, Massachussets Institute of Technology.

Patel, V. L., Frederiksen, C. H., & Groen, G. J. (1983). Relationship between comprehension and medical problem solving. In the essence of clinical competence: Psychological studies of expert reasoning in medicine. Symposium conducted at the Research in Medical Education Conference of the Association of American Medical Colleges, Washington, DC.

Patel, V. L., & Frederiksen, C. H. (1984). Cognitive processes in comprehension and knowledge acquisition by medical students and physicians. In H. Schmidt & H.G. DeVolder (Eds.), Tutorials in Problem Based Learning. Assen, Holland: Van Gorcum.

Patel, V. L., & Groen, G. J. (1986). Knowledge based solution strategies in medical reasoning. Cognitive Science, 10, 91-116.

Patel, V. L., Arocha, J. F., & Groen, G. J. (1986). Strategy selection and degree of expertise in medical reasoning. Paper in this volume.

Scardamalia, M., & Bereiter, C. (1984). Development of strategies in text processing. In H. Mandl, N. Stein, & T. Trabasso (Eds.), Learning and comprehension of text. Hillsdale, NJ: LEA.

Simon, D. P., & Simon, H. P. (1978). Individual differences in solving physics problems. In R. Siegler (Ed.), Children's thinking: What develops? Hillsdale, N.J.: Lawrance Earlbaum.