The Role of Knowledge in Teaching and Assessment of Problem-Solving

Geoffrey R. Norman
Associate Professor
Department of Clinical Epidemiology and Biostatistics
McMaster University
Hamilton, Ontario, L8N 3Z5

Recent research has resulted in a theory of medical problem-solving that is radically different from the theory which prevailed for nearly two decades. The new theory has not yet led to revisions in medical school curricula or evaluation methods, but potentially may engender change based upon a deeper understanding of the cognitive processes of expert clinicians. In this article, two theories of clinical problem-solving are contrasted. Studies of medical problem-solving conducted by the author and others are described. Evidence is presented to highlight the central role of experience in the development of competence. Finally, implications for the teaching and assessment of problem solving, especially in medicine, are suggested.

In the early 1960s medical education moved away from emphasizing the acquisition of rote knowledge towards the development of problem-solving strategies. Signaling this change were the "critical incident" study commissioned by the National Board of Medical Examiners, which clearly highlighted the multiple dimensions of clinical competence, and the appearance of objective measures of clinical problem-solving using card or paper based patient simulations (Rimoldi, 1961; McGuire and Babbott, 1967). These evaluation methods implicitly viewed "problem-solving" as a general skill or strategy applicable to different problems. Findings show moderate correlations between performance on these problems and measures of content knowledge were interpreted as evidence for the presence of a general problem-solving strategy (McGuire and Babbott, 1967), although equally low correlations were observed when scores on different problems were compared (Berner, 1977).

Curriculum efforts focusing on the development of problem-solving skills inevitably followed. Problem-based learning (Barrows and Tamblyn, 1981) in which basic science material was learned in the course of solving clinical problems, usually in a small group setting, exemplified the growing interest of educators in clinical problem-solving. The philosophy of this approach was simple—if students were to become effective problem-solvers, the faculty had a responsibility to teach problem-solving, and problem-based learning emerged as a natural approach to achieve this goal. However, since little was known about problem-solving, it was difficult to know what to teach. In response to this dilemma, studies were undertaken at McMaster University and Michigan State University in the early 1970s (Barrows et al., 1982; Neufeld et al., 1981; Elstein et al., 1978) to describe the problem-solving process of clinicians and its development in medical students. The results of these studies were inconsistent with the expectations of the investigators. They found a general process—the now classic "hypothetico-deductive method," but no evidence that some clinicians were consistently better or worse problem-solvers than others or that the problem-solving process improved from entry into medical school through clinical practice. Perhaps most disquieting was the finding that the best predictor of successful problem-solving, at least as assessed by obtaining the correct diagnosis, was the content of the diagnostic hypothesis.

As evidence of the fundamental role of early hypothesis generation, over 80% of the diagnostic hypotheses generated by clinicians occurred in the first quarter of the encounter. Of those encounters where the clinician arrived at the correct diagnoses 95% had thought of it within the first five minutes. Conversely, when physicians did not reach the correct diagnosis, 95% had not entertained it as an hypothesis. If clinicians thought of it in the first 5 minutes, they got it, otherwise they never did.

The conclusion from these data is that experts are experts because they think of better, more correct or more appropriate, hypotheses very early on. The search for a general problem-solving process supported the central role of knowledge, as it is reflected in the accurate hypotheses generated by experts.

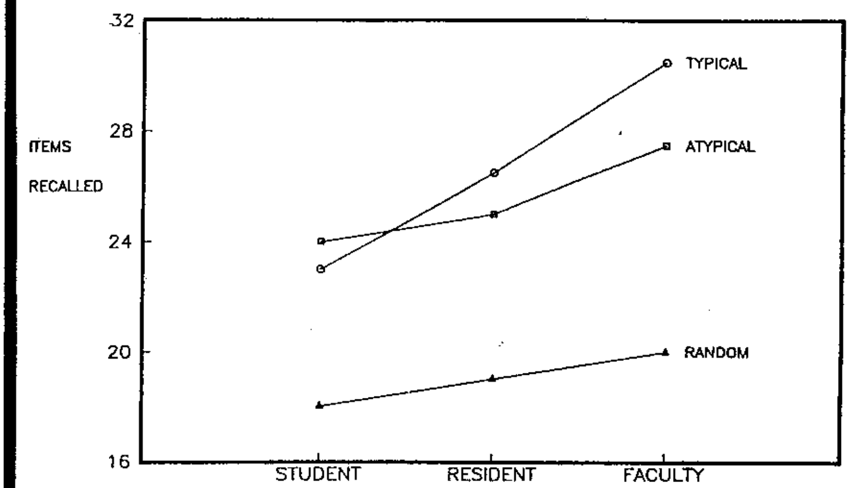
Studies in other domains, including chess (deGroot, 1965; Chase & Simon, 1973), the oriental game of GO (Reitman, 1976), physics (Larkin, et al. 1980), and other domains (Schoenfeld & Hermann, 1982), have arrived at similar conclusions. As an example, chess masters did not differ from novices in their thought processes—the number of moves considered, search strategies, depth of moves—but they thought of better moves (deGroot, 1965). As Glaser (1984) noted in his review article:

"Our interpretation is that the problem-solving difficulties of novices can be attributed largely to the inadequacies of their knowledge base and not to limitations in their problem-solving strategies." (p. 99)

One feature of memory which distinguishes experts from novices in chess is their ability to recall legal chess positions. After a five second viewing of a chess board, experts could remember better than 90% of pieces; however, if pieces were placed at random on the board no differences in recall between experts and novices were observed (deGroot, 1965).

These studies have been replicated in medicine (Norman et al., 1979). Four case protocols were developed, two typical cases of gall bladder disease and lung cancer and two "random" protocols, and were given to second year students, first and third year residents in family medicine, and family physicians. Each subject read the case through once at his own speed, and then recalled what was read. The amount recalled divided by time spent reading, for the typical and atypical protocols yielded results consistent with the chess findings. Typical cases showed a consistent increase in recall with increasing expertise, but no differences in recall were present with random protocols. One intriguing feature of this study was that the gradient of performance on typical cases began in residency education, suggesting that differences reflected a different kind of knowledge than the formal basic and clinical science knowledge taught in medical school.

Figure 1. Number of Items Recalled At Various Educational Levels.



Another intriguing finding emerged from a secondary analysis. One of the "random" protocols had been developed from an actual patient and might more correctly be called an "atypical" or unusual case. This seemed to be handled differently than the "true random" protocol. Items recalled for this case followed a gradient with education level similar to the typical cases (Figure 1). Time spent reading the protocol (Figure 2) was very different from either typical or random protocols, in that expert family physicians spent an inordinate amount of time puzzling over the case. This raises the issue of differences in expert behavior when dealing with typical and atypical cases, and what is it about a particular problem which makes it typical or atypical.

Typical Cases and Pattern Recognition

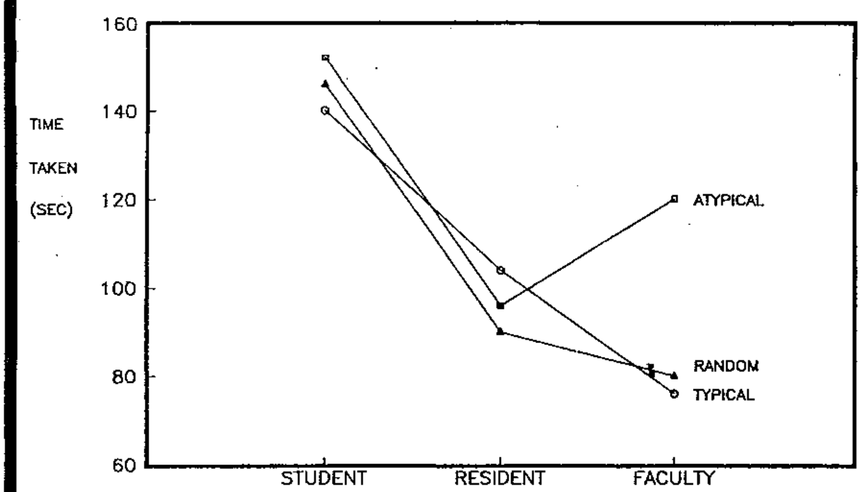
Typical cases probably represent 75 to 95% of the problems seen by a physician. One distinguishing feature of experts in dealing with these problems is their ability to see patterns or "chunks" of data, as evidenced by the natural grouping of phrases in free recall tasks (Figure 3). This can be measured by the time taken to recall an individual item, and studies of experts and novices in cardiology and rheumatology have demonstrated that this time is inversely related to expertise.

Two other phenomena occur in experts faced with this recall task. First, they tend to make large inductive inferences from the data, so that, for example, a productive cough, chronic and steadily worsening, over a period of several months or years, without hemoptysis, is labelled a "bronchitic cough." This appears to occur at a perceptual level so that experts are unable to recall individual features. Because of this phenomenon, studies of subspecialists could no longer demonstrate differences in total amount of recall with expertise (Muzzin et al., 1982, 1983); however experts could recall more important features—what Patel (1983) calls "frame" information, or what others have called "deep structure" (Chi et al., 1981). Finally, there is the rapidity or automaticity with which experts can appreciate patterns in data, best illustrated by visual tasks such as radiologic diagnosis. A study of chest film interpretation showed that experts were able to correctly identify abnormal films in an average of 7 seconds, and normal films in about 10 seconds (Norman, et al., 1984).

Automaticity is further supported by studies of disease prototypes (Bordage, 1984). A prototypical disease is one which best exemplifies the category, for example tuberculosis is a prototype of infectious pulmonary disease. Prototypes are not necessarily the most common or serious diseases within a category; rather, they are most representative in the sense that they have most features which distinguish that category from others. Disease prototypes are recalled earlier and more frequently by clinicians, are judged to have a greater number of overlapping features with the category, and are more rapidly and correctly classified as true or false when they are presented to clinicians in statements such as "diabetes is an example of an endocrine disease."

To summarize, experts process common disorders almost automatically at a perceptual level, so that the expert literally "sees" patterns which are not available to the novice. This processing, strongly influenced by prior experience and based on a pattern-matching process, shows little evidence of general problem-solving strategies which are independent of experience or knowledge, and may explain why studies of clinical problem solving have had difficulty revealing correlations in performance across problems.

Figure 2. Time to Read Protocol at Various Educational Levels.



Studies of Atypical Cases

As demonstrated in the earlier study of expert-novice memory, experts treated the single atypical case in the study very differently— recalling about as much data as in a typical case, but taking much more time. Similar results arose from two studies of visual processing—the radiology study, also discussed earlier, (Norman et al., 1985) and a study of recognition of skin lesions by students and clinicians (Norman et al., 1985). In both studies, correct interpretation of the stimuli by experts occurred more rapidly than by students. However, incorrect identification showed a positive relationship to expertise, so that expert dermatologists took an average of 25 seconds to incorrectly label a lesion, versus 14 seconds for clinical clerks and 11 seconds for first year students. It is significant that errors were associated with long processing times,— which reflect a much less automatic form of processing and probably more genuine "problem-solving." Chase and Simon (1973) arrived at a similar conclusion in studying chess masters.

None of these studies explicitly deals with what makes an atypical problem atypical. Bordage and Allen (1982) have, however, dealt explicitly with this issue by devising four types of written case protocols, the first corresponding to a textbook or prototypical presentation, and the remaining three containing either (a) a single critical cue necessary to support the correct diagnosis; (b) a single distracting cue unrelated to the diagnosis or (c) a misfitting cue which suggests an alternative, correct diagnosis. He showed that each atypical presentation leads to a large increase in error rates among both experts and novices, presumably because they demand a logical, cue-by-cue inference rather than pattern recognition. Placing these demands on the expert clinician results in a high error rate. Further research is required as to whether the differences in error rates and response times are in fact a reflection of qualitatively different processes, such as pattern recognition and problem-solving.

Implications for Education

Medical education in the twentieth century has been guided by two successive paradigms. The first rooted medicine firmly in the basic sciences and viewed clinical expertise as a mastery of a defined body of scientific facts and principles. The more recent paradigm has viewed the clinician as a problem-solver, distinguished by unique set of skills, but able to marshall and apply facts as needed to solve problems. To the extent that the interpretations made on the data here are correct, clinical expertise appears to be quite different from either characterization. Certainly, in viewing the diagnostic process, the distinguishing characteristic of the expert clinician is the availability of vast clinical experience against which new presenting situations can be matched in an almost intuitive and automatic manner, complemented by an occasional venture into a more psychologically or scientifically "pure" form of problem-solving behavior when the occasion arises. In this latter instance, one can presume that the clinician uses basic principles of anatomy, physiology and so on to arrive at a solution. No data have yet been collected on this aspect of problem-solving.

Any extrapolation from the limited data presented thus far to the educational domain must be viewed as conjectural. However, if clinical experience is a major determinant of expertise, it follows that the more clinical problems that a student can encounter, the better. Further, given the significant role of context in the ability to subsequently recall and recognize information (Riccio & Reitman, 1984), it is reasonable to suggest that since clinical and basic science "facts" will be used in a patient context, learning these principles in a patient situation should enhance recall. There is some evidence to support this conclusion from a study by Claessen and Boshuizen (1985) at Maastrict (a problem-based curriculum) and Utrecht (a "subject-based" curriculum). Using a recall task similar to those already described, there was consistent improvement in recall of the Maastrict students compared to Utrecht students as a function of educational level.

The implications for evaluation are also significant, if one accepts the theory that clinical problem-solving usually represents an interaction between the presenting situation and previous representations in memory. Since there is no reason to suppose that availability of appropriate representations for one problem would be related to a similar availability for a second, it may be that problem-solving is "content-specific," but not in the way it is commonly assessed on multiple choice tests. Rather, the "content" is more likely to represent complex, configural, and perhaps unconscious representations of signs and symptoms. This theory of expert memory would also account for the frequently observed lack of relationship between thoroughness in data gathering from patients experience. To the extent that experience leads to automatic associations between initial patterns and categories in memory, the expert will be seen to make apparently inductive "leaps" from minimal data, and the search for additional data is only confirmatory, or even irrelevant.

The study of general problem-solving strategies, if they exist, will require the careful selection of challenging problems similar to those of Bordage and Allen (1982). Under these circumstances, one should be able to observe the logical inference based on scientific principles which might be called "problem-solving." But in more common situations, these activities may well be absent.

Since the access of relevant knowledge in a patient context is the hallmark of expertise, the expectation that such knowledge should be available in other contexts, like multiple choice tests, stacks the cards against the expert. That increasing experience results in generally poorer performance on conventional tests is not surprising and should not be interpreted as an indicator of diminished clinical competence.

The studies discussed may be a first step towards the development of new testing methodologies. The methods are efficient, objective, and have shown high reliability across repeated cases—in contrast to many problem-solving tests. However, much work needs to be done before these approaches can be used as valid measures of competence. As yet, all that is certain is that scores are positively correlated with experience, and it remains to be demonstrated that they are related to competence.

Conclusions

Recent evidence from studies of the psychology of clinical reasoning suggests that expertise in medicine is characterized by the gradual accumulation of experience in memory which results in an ability to perceive, organize, and recall the extensive information presented by the patient. The experienced clinician is a better problem-solver by virtue of this accumulated experience, not as a result of any innate or learned problem-solving skill. These results suggest curriculum innovations focussed on the provision of an environment rich with clinical experience, and the learning of concepts in the context of clinical problems. Further research is required to identify the strategies used by experts in dealing with unusual or atypical problems.

References

Barrows, H.S., & Tamblyn, R.M. (1981). *Problem based learning: An approach to medical education*, New York: Springer Publishers.

Barrows, H.S., Norman, G.R.. Neufeld, V.R., & Feightner, J.W. (1982). The clinical reasoning of randomly selected physicians in general medical practice*/ Clinical and Investigative Medicine, 5,* 4955.

Berner, E.S., Bligh, T.J., & Guerin, R.O., (1977). An indication for a process dimension in medical problem-solving. *Medical Education, 11,* 324-328.

Bordage, G., & Allen, T. (1982). The etiology of diagnostic errors: *Process or content? Proceedings of the 21st conference on research in medical education,* Washington DC.

Chase, W.G., & Simon, H.A. (1973). Perception in chess. *Cognitive Science, 4,* 55-81.

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

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

De Groot, A. (1965). *Thought and choice in chess.* The Hague, Mouton.

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

Feltovitch, P.J. (1983). Expertise: Reorganizing and refining knowledge for use. *Professions Educational Research Newsletter, 4(3), 59.*

Glaser, R. (1984). Education and thinking: The role of knowledge, American Psychologist, 39, 93-103.

Larkin, J., McDermott, J., Simon, D.P., & Simon, H.A. (1980). Expert and novice performance in solving physics problems. *Science, 208,* 1335-42.

McGuire, C.H., & Babbott, D. (1967). Simulation technique in the measurement of problem solving skills. *Journal of Educational Measurement, 4,* 1-10.

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,* Washington, DC.

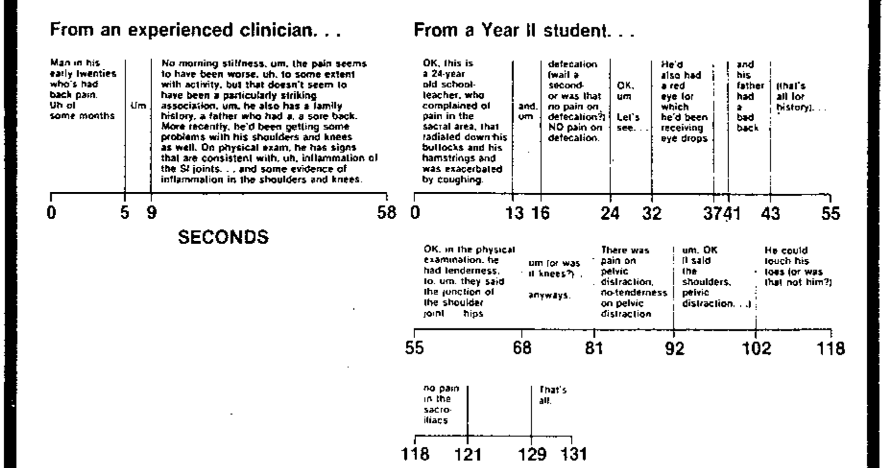
Muzzin, L.J., Norman, G.R., Feightner, J.W., & Tugwell, P. (1983). Expertise in recall of clinical protocols in two specialty areas. *Proceedings of the 22nd conference on research in medical education,* Washington, DC.

Neufeld, V.R., Norman, G.R., Barrows, H.S., & Feightner, J.W. (1981). Clinical reasoning of medical students: A cross-sectional and longitudinal analysis. *Medical Education, 15,* 315-327.

Norman, G.R., Muzzin, L.J., & Rosenthal, D. (1985). Expert-novice differences in perception and categorization in dermatology. *American Educational Research Association Annual Meeting,* Chicago, IL.

Norman, G.R. (1984). Cognitive psychology and medical problem-solving: Memory studies. *American Educational Research Association Meeting,* New Orleans, LA.

Figure 3. Examples of Transcribed Recall from an Experienced Clinician and from a Year II Student.



	EXPERT	NOVICE
No. of items Recalled	12/28	14/28
No. of "runs"	2	9
Avg. length of runs	27 sec.	9 sec.
Avg. pause length	4 sec.	7.5 sec.
Items/Run	6	1.6

Patel, V.L. (1984). Expert diagnostic reasoning: An interactive process of case comprehension. *Professions Educational Research Newsletter, 5*(3), 7-11.

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

Riccio, D.C., & Richardson, R. (1984). Memory retrieval deficits based on altered contextual cues: A paradox. *Psychological Bulletin, 96*, 152-165.

Rimoldi, H.A. (1961). The test of diagnostic skills, *Journal of Medical Education, 36*, 73-79.

Schoenfeld, A.H., & Hermann, D.J. (1982). Problem perception and knowledge structure in expert and novice mathematical problem solvers. *Journal of Experimental Psychology: Learning, Memory and Cognition, 8*, 484-494.