Medical Education 1990, **24,** 396–399

SHORT COMMUNICATION

Clinical inquiry and scientific inquiry

G. R. NORMAN, V. L. PATEL† & H. G. SCHMIDT‡

Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, †Centre for Medical Education, McGill University, Montreal and ‡University of Limburg, Maastricht

Key words: *clinical competence; *diagnosis; *science; problem-solving; research; education, medical

In a recent position paper (Barrows 1990), Barrows described an important component of the clinical reasoning process which he calls 'inquiry' — the act of eliciting data from the patient by history and physical examination. In his words 'An effective inquiry is tuned to the likely diagnostic possibilities suggested by the patient's problem as it unfolds, asking those questions and performing those items of examination that will help assemble the information needed to support the more likely diagnosis(es) and separate them from others entertained'. He further claims that these actions are a 'logical, deductive, reasoning skill'.

It would be heretical to dispute the value of such a cherished ideal as the effective, efficient history and physical examination. Nevertheless, Barrows' extrapolation from a defence of history and physical examination skills to a condemnation of particular educational and research methods in which the learner or subject is provided with a data base and does not inquire actively for patient data moves from the realm of clinical inquiry to educational and scientific inquiry, areas where we believe scientific evidence may be brought to bear.

Correspondence: Dr G. R. Norman, Department of Clinical Epidemiology and Biostatistics, McMaster University Health Sciences Centre, 1200 Main Street West, Hamilton, Ontario, Canada L8N 3Z5.

In the spirit of scientific inquiry then, we may begin by examining the evidence in support of the premise that this inquiry process is 'an important and central task for clinicians in their daily work with patients'. Interestingly, Barrows' own work provides some disconfirming evidence (Neufeld *et al.* 1981; Barrows *et al.* 1982.) The studies examined the activities of students — first, second and third year — and clinicians, family doctors and consulting internists, as they worked up clinical problems presented by simulated patients. Data gathering, or 'inquiry', was characterized by about 20 different variables including the number of history questions, the number of physical examination manoeuvres, the proportion of time in non-routine or routine questions, the number of hypothesis-testing questions, the number and proportion of significant findings elicited, the weight of findings against hypotheses, etc.

As a measure of expertise, these were suspect since there was no consistent improvement in any of these measures from undergraduate to doctor. Considering the generalizability of 'inquiry skills' across encounters, 16 correlations were examined; there was a mean of 0·05 and two significant negative and two significant positive correlations. Finally, neither the number of history questions, the number of physical examinations, the thoroughness of data gathering, nor any of the measures related to 'inquiry' predicted diagnosis or management outcomes.

A more recent study (Norman *et al.* 1986) used standardized patients and students from all educational levels, each of whom saw a total of 10 clinical problems from two subspecialty areas. Again, however, the simple correlation between various measures of data gathered and diagnosis was small (about 0·25), and the partial correlation controlling for educational level was only 0·12–0·15 and not significant. Thus the acquisition of good 'inquiry skills' is not generalizable, is not related to expertise, and is not related to diagnostic and management outcome in the studies reviewed. Of course the measures may be at fault, although it is difficult to contemplate other ways in which data-gathering may be operationalized.

Thus, although 'inquiry' may be a central skill, it is not easy to demonstrate its centrality empirically. Of course, whatever the evidence, few clinicians or educators would favour abandoning the teaching of history and physical examination skills in medical school (and we would not be among them). But that is not Barrows' central point: instead he is calling for a critical examination of those who, in the development of materials for problem-based learning (PBL), provide the data base, thereby denying students the opportunity to practise inquiry skills. It would seem self-evident that providing the data would be an inferior teaching strategy, if the alternative were feasible. Still, to claim that this would be 'worse than spoon-feeding' in terms of its effect on student performance is a bit strident. Most studies of the impact of PBL, compared to drastically different teaching approaches, have demonstrated little or no difference (Schmidt *et al.* 1987a), and it would appear unlikely that such studies would be able to demonstrate the impact of a relatively small curriculum innovation like the format of the written problems. Still, we might note that Barrows' viewpoint is consistent with one model of PBL, that the goal of PBL is to acquire problem-solving skills. However, no study to date is available showing that problem-based curricula indeed facilitate the acquisition of these skills.

For the most part, programmes aimed at teaching problem-solving skills generally have been unsuccessful (for a critical discussion see Perkins & Salomon (1989)). It may come as no surprise, then, that some authors (e.g Schmidt 1982; Norman 1988) have challenged Barrows' view of PBL as a vehicle for teaching inquiry skills, and have suggested that PBL most and for all is a useful stategy for the acquisition of knowledge organized around problems. There is some evidence that this is precisely what PBL does (Schmidt *et al.* 1989). In this latter perspective, however, the precise sequence of the data is probably of less importance.

What then of the *research* strategy which involves presenting the subject (student) with a patient data base, and asking for elaboration, discussion or recall of the clinical material? As researchers who have used this method for over a decade with, in our view, considerable success (see for example Schmidt *et al.* 1976; Patel *et al.* 1988; Norman *et al.* 1989), we would be distressed that 'Putting clinicians or student subjects through a patient problem-solving experience in which they cannot inquire . . . runs the risk of producing artificial results' and that 'Any research aimed at elucidating information about the clinical reasoning process that does not allow for the central skill of inquiry to occur must remain suspect'. Although we would be the first to agree that, other things being equal, a close approximation to reality is a desirable goal, we believe the adequacy of a scientific experiment should not be judged simply on grounds of closeness to reality, but must consider the theoretical basis and experimental rigour of the method. If biomedical researchers were forced by convention to restrict their observations to those gathered on living human species by the argument that *Escherichia coli* or mice do not look like or act like real persons, there would be little for biological scientists to talk about. Few would argue that biomedical researchers do not make a valid contribution to science, using measures of lower fidelity that allow for more rigorous experimental control. We would contend that so-called 'artificial' experimental tasks can similarly reveal much about the process of medical problem-solving.

From the outset, we should note that these investigations are explicitly *not* directed at the observation of a problem-solving process — ample evidence exists that this line of inquiry has not been particularly fruitful (Elstein *et al.* 1978; Norman 1988). In fact, a recent paper (Norman 1990) showed that measures based on recall of complete laboratory data bases were demonstrably more capable of discriminating between levels of expertise and between individuals within a level of expertise than measures based on actual performance using simulated patients or conventional written tests (Norman 1990). Instead, these research methods are directed to an understanding of the expert's mental organization of knowledge and its relation to expertise.

Students are asked to engage in a variety of tasks — sequentially representing the patient's problem as additional data are provided (Feltovich *et al.* 1984; Joseph & Patel 1990), explaining the problem in terms of basic science concepts (Patel & Groen 1986), solving the problem and then incidentally recalling specific data (Schmidt *et al.* 1988; Norman *et al.* 1989a), or rapidly recognizing visual patterns on radiographs or photographs (Lesgold *et al.* 1988, Norman *et al.* 1989b). These methods allow us to describe more precisely various aspects of diagnostic thinking such as the nature of knowledge structures, the use of problem-solving strategies, and hypothesis generation along a developmental continuum from novice student to expert doctor. These processes are not easily amenable to analysis using high fidelity patient simulations.

This theoretical approach has been adapted from methods used by researchers attempting to understand the evolution of expertise in domains ranging from chess (Chase & Simon 1973) to computers (Adelson 1987), and dates over several decades to studies of chess masters. De Groot (1963) demonstrated that the most sensitive measure of expertise in chess was based on simple recall of a mid-game chess position. Although De Groot was himself a master, and spent many hours observing other masters at play, in the end this measure of recall was a better probe for understanding the nature of expertise in chess than observation or thinking aloud. We might point out in passing that the mid-game positions were provided as a complete data base to the subjects — they did not get to the position of their own accord by a process of chess 'inquiry'.

If, then, these methods show such promise in understanding the complexities of clinical reasoning, what insight do they reveal about the problem which emerged earlier in this paper? Why, if inquiry skills are so central and have endured so well the ravages of time, are they so apparently superfluous when subjected to empirical scrutiny? There are a number of possibilities. First, the perception studies of Allen *et al.* (1988) demonstrated that the initial diagnostic hypothesis not only emerges early in the encounter, but also is strongly influenced by similarity to prior cases, rather than by the results of systematic inquiry for individual features. Although derived from dermatology, a highly visual domain, similar results have been reported in internal medicine (Van Rossum *et al.* 1989). Once a hypothesis is entertained, inquiry is primarily focused on marshalling support for the hypothesis, and the more expert the clinician, the stronger the search for confirmatory information (Barrows *et al.* 1978). Features may actually be reinterpreted to coincide with the preferred hypothesis (Allen *et al.* 1988). Despite the claim of Barrows that 'efficient and effective problem-related inquiry skills require a higher order of cognitive skills based on a good understanding of disease pathophysiology', evidence from this research is that experts only rarely use pathophysiological concepts in their reasoning (Kaufman & Patel 1987; Bushuizen *et al.* 1989), not because they have lost this knowledge but because it is unnecessary for their solutions, except in rare circumstances (Boshuizen *et al.* 1989). Finally, because experts see patterns faster (Norman *et al.* 1989b), they are apparently able to come to closure with less information than novices.

In short, before dismissing any research methodology, it is well to explore the explanatory power of the paradigm: the evidence that the method can provide useful explanations for theoretical and practical problems. In this case, we will let the evidence speak for itself.

References

Adelson B. (1987) When novices surpass experts: the difficulty of a task may increase with expertise. *Journal of Experimental Psychology: Learning, Memory and Cognition* **10**, 483–95.

Allen S. W., Brooks L. R. & Norman G. R. (1988) Effect of prior examples on rule-based diagnostic performance. In: *Proceedings of the 27th Conference on Research in Medical Education Chicago, Illinois*, pp. 9–14. Association of American Medical Colleges, Washington, DC.

Barrows H. S., Feightner J. W., Neufeld V. R., & Norman G. R. (1978) *An Analysis of the Clinical Methods of Medical Students and Clinicians*. Final Report to the Ontario Ministry of Health. McMaster University, Hamilton, Ontario.

Barrows H. S. (1990) Inquiry: the pedagogical importance of a skill central to clinical practice. *Medical Education* **24,** 3–5.

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,** 49–53.

Boshuizen H. P. A., Schmidt H. G. & Coughlin L. D. (1989) On the application of medical basic science knowledge in clinical reasoning: implications for structural knowledge differences between experts and novices. In: *Proceedings of the 10th Annual Conference of the Cognitive Science Society,* pp. 517–23. Erlbaum, Hillsdale, New Jersey.

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

de Groot A. (1963) *Thought and Choice in Chess.* Mouton, The Hague.

Elstein A.S., Shulman L. S. & Sprafka S. A. (1978) *Medical Problem-Solving: An Analysis of Clinical Reasoning.* Harvard University Press, Cambridge, Massachusetts.

Feltovich P. J., Johnson P. E., Moler J. H. & Swanson D. B. (1984) LCS: The role and devlopment of medical knowledge in diagnostic expertise. In: *Readings in Medical Artficial Intelligence: The First Decade* (ed. by W. J. Clancy & E. H. Shortcliffe), pp. 275–31. Addison-Wesley, New York.

Kaufman D. R. & Patel V. L. (1987) Causal reasoning in medical explanations: Expert – intermediate – novice differences. Paper presented at the American Educational Research Meeting, Washington, 1987.

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

Lesgold A., Rubinson H., Feltovich P., Glaser R., Klopfer D. & Wang Y. (1988) Expertise in a complex skill: diagnosing x-ray pictures. In: *The Nature of Expertise* (ed. by M. Chi, R. Glaser & M. Farr), pp. 311–42. Lawrence Erlbaum, Hillsdale, New Jersey.

Neufeld V. R., Norman G. R., Barrows H. S. & Feightner J. W. (1981) Clinical problem-solving by medical students: a cross-sectional and longitudinal analysis. *Medical Education* **15,** 315–22.

Norman G. R., Tugwell P., Feightner J. W., Muzzin L. J. & Jacoby L. L. (1986). Knowledge and clinical problem solving. *Medical Education* **19,** 344–56.

Norman G. R., Brooks L. R. & Allen S. W. (1989a) Recall by expert medical practitioners and novices as a record of processing attention. *Journal of Experimental Psychology: Learning, Memory and Cognition* **15,** 116–74.

Norman G. R., Brooks L. R., Allen S. W. & Rosenthal D. R. (1989b) The development of expertise in dermatology. *Archives of Dermatology* **125,** 1062–5.

Norman G. R. (1988) Problem-solving skills, solving problems, and problem-based learning. *Medical Education* **22,** 279–86.

Norman G. R. (1990) Reliability and validity of some cognitive measures of clinical reasoning. *Teaching and Learning in Medicine* **1,** 194–200.

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

Patel V. L., Evans D. A. & Groen G. J. (1988) Biomedical knowledge and clinical reasoning. In: *Cognitive Science in Medicine* (ed. by V. L. Patel & D. A. Evans), pp. 53–112. MIT (Bradford) Press, Cambridge, Massachusetts.

Perkins D. N. & Salomon G. (1989) Are thinking skills context-bound? *Educational Researcher* **18,** 16–25.

Schmidt H. G. (1982) *Activatie van voorkennis, intrinsieke motivatie en de verwerking van tekst, studies in problemgestuurd onderwijs.* (Activation of prior knowledge, intrinsic motivation and text processing, studies in problem-based learning). Doctoral Dissertation. Van Walraven, Apeldoorn.

Schmidt H. G., Boshuizen H. P. A. & Hobus P. P. M. (1988) Transitory stages in the development of medical expertise: The "intermediate effect" in clinical case representation studies. In: *Proceedings of the Tenth Annual Conference of the Cognitive Science Society, Montreal Canada,* pp. 139–45. Erlbaum, Hillsdale, New Jersey.

Schmidt H. G., Dauphinee D. G. & Patel V. L. (1987a) Comparing the effects of problem-based and conventional curricula in an international sample. *Journal of Medical Education* **62,** 305–15.

Schmidt H. G., Hobus P. P. M., Patel V. L. & Boshuizen H. P. A. (1987b) Contextual factors in the activation of first hypotheses: Expert-novice differences. Paper presented at the Annual Meeting of the American Educational Research Association. Washington, DC.

Schmidt H. G., De Grave W. S., De Volder M. L., Moust J. H. C. & Patel V. L (1989) Explanatory models in the processing of science text: the role of prior knowledge activation through small-group discussion. *Journal of Educational Psychology* **81,** 481–91.

Van Rossum H. J. M., Briet E., Bender W. & Meenders A. E. (1989) The transfer effect of a single patient demonstration on diagnostic judgement at graduate medical students: both better and worse. *Proceedings of the Third International Conference on Teaching and Assessment of Competence, Groningen, The Netherlands, 1989.*

Received 28 *February* 1990; *accepted for publication* 5 *March* 1990