Research in clinical reasoning: past history and current trends

Geoffrey Norman

BACKGROUND Research in clinical reasoning has been conducted for over 30 years. Throughout this time there have been a number of identifiable trends in methodology and theory.

PURPOSE This paper identifies three broad research traditions, ordered chronologically, are: (a) attempts to understand reasoning as a general skill – the 'clinical reasoning' process; (b) research based on probes of memory – reasoning related to the amount of knowledge and memory; and (c) research related to different kinds of mental representations – semantic qualifiers, scripts, schemas and exemplars.

RESULTS AND CONCLUSIONS Several broad themes emerge from this review. First, there is little evidence that reasoning can be characterised in terms of general process variables. Secondly, it is evident that expertise is associated, not with a single basic representation but with multiple coordinated representations in memory, from causal mechanisms to prior examples. Different representations may be utilised in different circumstances, but little is known about the characteristics of a particular situation that led to a change in strategy.

IMPLICATIONS It becomes evident that expertise lies in the availability of multiple representations of knowledge. Perhaps the most critical aspect of learning is not the acquisition of a particular strategy or skill, nor is it the availability of a particular kind of knowledge. Rather, the critical element may be deliberate practice with multiple examples which, on the hand, facilitates the availability of concepts and conceptual knowledge (i.e. transfer) and, on the other hand, adds to a storehouse of already solved problems.

Michael De Groote Centre for Learning and Discovery, McMaster University, Hamilton, ON, Canada

Correspondence: Geoffrey Norman PhD, Assistant Dean, Program for Educational Research & Development, Michael De Groote Centre for Learning and Discovery, Room 3019, McMaster University, 1200 Main St.W., Hamilton, ON, L8N 3Z5, Canada. Tel: (905) 525 9140, ext.22119; Fax: (905) 572 7099; E-mail: norman@mcmaster.ca

KEYWORDS education, medical, undergraduate/*methods; clinical competence/*standards; problem solving; decision making; research design/*standards; teaching/methods.

Medical Education 2005; ***39****: 418–427*
doi:10.1111/j.1365-2929.2005.02127.x

INTRODUCTION

Educators agree that clinical reasoning is a central component of physician competence, and objectives related to mastery of clinical reasoning skills appear in the documentation of most medical schools, licensing bodies and speciality societies. All can agree that clinical reasoning, or one of its many synonyms – problem-solving, decision-making, judgement – should be taught and tested. But once one goes beyond the phrase to attempt to determine what it is, or to devise instructional approaches or testing methods, matters become much more complicated.

Research attempting to understand the nature of clinical reasoning has been under way for nearly 3 decades. Although the number of investigators contributing to the field is limited, progress has been substantial, and we now have a much better understanding of factors related to expertise in clinical medicine. Regrettably, this literature is difficult to access and synthesise for 2 reasons: first, the studies are published in areas as diverse as the perspectives of the researchers themselves. In particular, while many studies appear in medical education journals, others are scattered through areas such as sociology, cognitive psychology and clinical psychology. Secondly, it

 © Blackwell Publishing Ltd

Overview

What is already known on this subject

Clinical reasoning has been a fruitful area of research for 3 decades. However, theories and evidence are poorly synthesised and interpretation is difficult.

What this study adds

Research is summarised into broad historical domains – expert reasoning as a process, as memory, as knowledge representations. The study synthesises and critiques the multiple approaches to research on knowledge representation. Implications for teaching and assessment are highlighted.

Suggestions for future research

Instead of attempting to determine the one 'best' representation of knowledge, researchers should recognise that experts access multiple representations and should explore the conditions under which these are used.

The importance of deliberate practice in expertise has not been addressed adequately.

frequently appears that there is little consensus among investigators about some of the most basic characteristics of clinical reasoning.

In this review, I will attempt to synthesise this heterogeneous literature. The summary will necessarily be more descriptive of various approaches and findings than prescriptive. However, emerging from this synthesis will be some fundamental questions and issues directed at both researchers and practitioners.

While it is easy to subsume almost all of what doctors do (with some exceptions such as interpersonal skills and technical skills) under the term 'clinical reasoning', practically all the research has a more limited purview, focusing instead on the processes doctors use to arrive at an initial diagnosis based on history and physical examination (and occasionally investigations). While the limitations of this perspective have been acknowledged for a long time,[1] diagnosis remains a central focus for research. Decision analysis is an exception to this rule; it is concerned primarily with appropriateness of management decisions under uncertainty. However, the methods of decision analysis are directed more at what doctors *should* do (using formal mathematical methods based on probability and utility of various possible outcomes) than what they *do* do. Consequently, I will not discuss these approaches further.

The review is designed to address two primary questions:

1. To what extent is clinical reasoning a general skill; to what extent is it a consequence of the application of specific knowledge? Can it be taught as a general skill? Can we assess reasoning skill independent of knowledge?
2. What kind of knowledge is necessary for successful clinical reasoning? Are some kinds of knowledge more fundamental than others? Does the kind of essential knowledge vary by level of expertise? By specialty or discipline? By context (e.g. primary care vs. tertiary care)?

While the framework for the review is historical, the answers to these two questions underlie all the work, and emerge naturally from the accumulation of evidence.

EARLY RESEARCH ON CLINICAL REASONING: THE GOLDEN AGE OF PROBLEM-SOLVING

In the early 1970s 2 research groups, at Michigan State University and McMaster University, began observational studies directed at understanding clinical problem solving. Experienced clinicians and students at various levels were observed with standardised patients, and encouraged to 'think aloud' (Michigan State) or subsequently reviewed a videotape of their interactions, as a 'stimulated recall' of their thought processes (McMaster). A general model of clinical problem-solving, the 'hypothetico-deductive' method, emerged from both studies. Within a few minutes of the beginning of the encounter clinicians generated several diagnostic hypotheses, and gathered subsequent data to rule in or out these hypotheses. The only problem is that the process was too general; all subjects at all levels were doing about the same thing.[2,3] What distinguished experts from novices was that they generated not more, not quicker, but better hypotheses, and the accuracy of the early hypotheses was a strong predictor of their final conclusion.[4] Further, success on one problem was a poor predictor of success on a second problem; typical correlations across problems were of the order of 0.1–0.3, a finding labelled 'content specificity' by Elstein.[2]

Both findings cast a long shadow over the notion of a general problem-solving process or skill associated with clinical expertise or experience. On one hand, the process was too general and unrelated to expertise. On the other hand, the outcome, such as diagnostic accuracy, was apparently related strongly to content knowledge, not to a general process. The direct consequence was that these early studies were not followed up and a new generation of researchers changed direction, towards a focus on expert knowledge, and away from an expert process. However, as we shall see, the break was not quite a clean as it may seem.

THE 1980S: THE AGE OF MEMORY

With the failure of general reasoning skills, the research community searched for other explanations. One arose from the growing literature on expertise in other domains, particularly chess. It emerged that the single best measure of expertise in chess was recall of a typical mid-game position, where after a 5 second exposure experts would typically recall the exact position of about 80% of the pieces. Underlying the relation between memory performance and expertise was the notion that expertise was, to a large degree, a matter of acquiring a large set of representative cases which can be used for analogies to a new problem situation. It is estimated that a chess master has remembered 50 000 game positions.[5]

Because this appeared to be a good measure of expertise in a number of domains, and because it responded to the idea that expertise was related more closely to specific knowledge than general skills, a number of investigators attempted to replicate the findings in medicine. However, failures[6–8] far exceeded successes.[9] One explanation that has been carefully explored by Schmidt[10] is that the expert has access to extensive case knowledge, but this knowledge remains 'encapsulated' until needed. Because relatively common cases contain considerable redundancy, this encapsulated knowledge does not emerge with the usual recall tasks. Expertise effects do, however, emerge when the task demands are increased, either by reducing time[10] or increasing case complexity.[9,11]

In hindsight, it is likely that the failure to show a relation between expertise in medicine and recall is a consequence of an incomplete understanding of the diagnostic task. In chess, keeping track of every piece is critical to success.[12] In medicine, there is little gain from gathering and remembering extensive amounts of patient data, consequently thoroughness is a poor index of expertise.[3,13] However, this may not be true in all clinical domains and all circumstances. In nephrology, where expertise resides in part in unravelling complex and non-linear relationships among laboratory data, some studies have shown a positive relation between memory for numerical laboratory data and expertise.[9,11] Nevertheless, the findings appear to indicate that, unlike chess and some other domains, most kinds of medical expertise may not be particularly related to the ability to recall patient data.

A larger question is why memory recall is such a good measure of expertise in chess and elsewhere, but not medicine. Ericsson[14] and others have shown that expertise in chess derives primarily from practice with thousands of cases, and the memory recall studies are probably tapping into recall of cases, hence are related strongly to expertise. There is every reason to presume that clinical expertise is also dependent on extensive practice, but memory of features is not capturing this retrieval mechanism. Regardless of the explanation, the failure of memory recall methods then led to a second revolution within the field. Because total amount of knowledge was not informative, perhaps expertise related instead to the kinds of knowledge the expert had, and the ways that the knowledge was organised. The consequence was a proliferation of studies directed at knowledge representations.

THE 1990S: THE AGE OF MENTAL REPRESENTATIONS

While memory recall appears to be a poor measure of expertise, nevertheless it seems almost axiomatic that experts must have available to them more knowledge, of more kinds, which is better organised and more accessible, than novices. If so, then careful study of knowledge organisation should yield valuable insights into the nature or reasoning.

However, before we enter this review, a brief methodological note is in order. Experts in any domain have more knowledge of many kinds than novices. In medicine, expert clinicians probably have more knowledge and better understanding of the mechanisms underlying disease (although this is not always apparent on standardised tests,[15]), more knowledge of the clinical manifestations of rare diseases, better intuitions about probabilities and base rates and much more experience with individual cases. A reasonable conjecture is that they know more about everything in their field than novices. Consequently, showing that experts have more knowledge of a particular type (such as probabilities) than novices is necessary, but not sufficient to claim that this knowledge is central to expertise. Similarly, showing that a type of knowledge is correlated with expertise or success is not sufficient to argue that it is central to expertise, as both knowledge and success may be associated with some other variable. People who are able to solve a problem probably have more knowledge about the problem, of all kinds, than those who are unsuccessful.

In a seminal paper, Schmidt *et al.*[16] suggested a developmental theory of medical expertise which explicitly recognised 3 distinctly different kinds of knowledge associated with expertise and posited that, with increasing experience, clinicians move through 3 kinds of mental representations, from basic mechanisms of disease to illness scripts to exemplars derived from experience. Of course, this reflects educational experiences: emphasis on mechanisms in the pre-clinical years, learning of rules of diagnosis as junior clinicians, supplemented further by case experience, in the latter undergraduate years, in residency, and beyond. What was alluded to, but left unanswered in that paper, was the relative contribution of the various knowledge forms to expert performance, and whether each representation is available and used depending on the particular context. It was conjectured that with difficult problems the expert might invoke more basic kinds of knowledge, but no evidence was available.

Research during the 1990s focused on these forms of mental representations, and can be divided roughly into the broad categories considered by Schmidt *et al.*[16] basic science or causal knowledge, schema, scripts and other representations of the relation between signs and symptoms and diagnoses, and exemplars based on experience with past cases.

The role of basic science

Given the amount of educational time devoted to learning basic science, it is reasonable to assume that expert clinicians exemplify the application of scientific concepts to clinical problems. Such an assumption guided 2 major research programmes in the 1980s, by Henk Schmidt[17] and Vimla Patel.[18] The basic methodological strategy was similar: the experimenter provided written case scenarios and encouraged participants to discuss their reasoning, then looked for instances of application of science knowledge, or alternatively provided both a basic science test and a clinical case and looked for integration of the basic science concepts into the solution of the clinical case.

The resulting verbal 'think-aloud' protocols were then analysed using propositional analysis methods. Patel and Groen[7] showed that experts had more coherent explanations for the problem, were more selective in the use of data and made more inferences from the data. However, surprisingly, experts used less basic science in their explanations than relative novices. Schmidt[17] has used similar methods of proposition analysis based on think-aloud protocols, and arrived at similar conclusions, primarily that expert clinicians make little use of biomedical science in daily reasoning. However, their investigations have gone further. To explain why experts mention less basic science and also to solve the problem of reduced recall of experts in memory studies, they postulate that this knowledge is encapsulated, as discussed earlier.[10]

One reason that experts may not use basic science is that the strategy of teaching in traditional courses encourages a separation between basic science and clinical, which are never reintegrated. There is some evidence of this; Patel *et al.*[19] showed that students in a traditional programme did not integrate basic science concepts into solutions of clinical cases, but students from a problem-based learning (PBL) programme did. However, PBL students were less accurate in diagnosis and made more conceptual errors. On the other hand, studies by Schmidt's group[20] showed more integration by PBL students but fewer conceptual errors, so this may be a local effect.

However, it seems unlikely that curriculum effects in the first 2 years of medical school would persist over a career. An alternative explanation is simply that most of the time, in most disciplines, experts do not need basic science. (If we have memorised the solution of 14×14, we do not need to remember the rules of multiplication to solve it.) However, in difficult or ambiguous situations, experts may revert to basic science explanations which, as Schmidt[8] has shown, are available but usually not accessed. For example, Norman[20] showed that with very difficult nephrology cases, expert nephrologists used far more basic science propositions than novices and were much more successful.

Thus it appears that, while expert clinicians may have superior knowledge of the relevant basic science concepts in their area, they rarely make use of this knowledge and mobilise it only to solve difficult or complex problems. Of course, the research is hardly exhaustive; conceptual knowledge may well be used extensively in management decision-making and in disciplines such as anesthesiology or intensive-care medicine, which have not been examined.

Nevertheless, that does not really explain the extensive time devoted to basic science instruction in most medical schools. One possibility is that an understanding of mechanisms provides some coherence to the lists of the signs and symptoms of various diseases that must be learned in the clinical years. Woods[21] has shown that participants who learn a 'basic science' explanation of a disease actually improved their diagnostic accuracy on representative cases after some time delay; whereas performance of a comparable group who simply memorised the features fell off after a 1-week delay.

Formal knowledge of diseases probabilities, scripts, schemas

Basic science is only one kind of knowledge in clinical medicine. Students also must spend many hours learning 'the 29 causes of anaemia', 'the signs and symptoms of ALS' and 'the differential diagnosis of weakness', the formal knowledge of clinical medicine. Indeed, when clinicians and students discuss diagnosis, it is this kind of knowledge which is central. An investigation of expertise must therefore consider how this knowledge is represented in the mind.

Perhaps the simplest kind of representation would be a list-like structure 'the ''signs and symptoms of'''. In the literature, this representation is perhaps allied most closely with the idea of scripts or schemas, borrowed from cognitive psychology.[22] Barrows and Feltovich[23] proposed the idea of 'illness scripts' as a story-like narration of a typical case of the condition. Building on this idea, Charlin[24] has developed an assessment method in which the student must complete a matrix of signs and symptoms against hypotheses, inserting appropriate weights. A further extension of this idea is the claim that expert clinicians have mental probability matrices (e.g. the probability of diaphoresis with myocardial infarction)[25] and that their reasoning amounts primarily to a Bayesian combination of probabilities. Another kind of schema is proposed by Mandin,[26,27] involving decision trees beginning with a clinical presentation (e.g. fatigue, hyponatraemia) and ending with specific diagnoses. One study showed that both experts and novices use 3 approaches to solving problems – hypothetico-deductive, pattern recognition from experiential knowledge, and schema induction, and that the latter 2 were associated more strongly with diagnostic accuracy than the first.[27] The theory has served as the basis for an entire medical school curriculum based on 126 presentations.[28]

If experts use scripts or schemas to represent diseases, how do expert scripts differ from novices? Bordage[29,30] has studied extensively the ways that novices and experts describe the features of diseases, examining the extent to which participants use what he calls 'semantic qualifiers' (SQs) – standard representations, usually bipolar, of signs and symptoms (such as proximal vs. distal, large joint vs. small joint, recurrent vs. acute or chronic). In turn, he has characterised different levels of expertise related to how these SQs are organised: 'reduced' – few features with no linkages, 'dispersed' – extensive but disorganised, 'elaborated' – extensive use of SQs with clear associations, and 'compiled – rapid and correct summary. These levels have been shown to be related to diagnostic accuracy.

The reader may well feel a sense of frustration at this point. We have presented evidence that expertise is distinguished by acquisition of illness scripts, decision trees, symptom × disease probabilities, semantic qualifiers and more (or less) basic science. Surely all these models of mental representations cannot be right? The answer is a qualified 'yes and no'. It is indeed possible that expertise in a domain results in an accumulation of knowledge of all kinds; indeed it would seem that it could hardly be otherwise. But there remains another kind of knowledge to be considered.

Experiential knowledge: exemplars

So far, we have examined evidence for the existence of various forms of analytical or formal knowledge. Many are variants of the kind of knowledge found in textbooks of physiology or clinical medicine. Yet it is clear on reflection that this form of knowledge is not all there about expertise. It takes many years of clinical practice before the student is deemed to have acquired enough experiential knowledge to be fit for independent practice. We have not yet considered the role of experiential knowledge in clinical expertise.

Norman's research programme has focused on this kind of knowledge, using a framework derived from exemplar models of categorisation in psychology.[31,32] The basic idea is that many categories we use in our representations of the world are defined, in part, by a large collection of examples derived from past experience (the categories 'dog' or 'chair' contain many specific examples in memory), and when we must classify an object we do it by rapid retrieval of a similar prior example, without conscious awareness.

In order to show the influence of a specific prior exemplar, the studies typically have 2 phases; a learning phase where specific examples are learned (and experimentally manipulated) and a test phase, where the influence of these examples on test cases is explored. Large effects have been shown in dermatology, where similarity to a prior case results in gains of accuracy of about 40% with residents[33] and 28–44% with medical students.[34] In these studies similarity was based on highly visual materials, so it may not generalise. For that reason, another study was conducted on electrocardiographic (ECG) diagnosis[35] and the manipulation was actually conducted on verbal information, the age, gender and occupation of the patients, which was objectively irrelevant to the diagnosis. The test materials were designed to be ambiguous with 2 probable diagnoses; however, if the non-relevant historical information were used to recall a prior example in the series, this would lead to the incorrect diagnosis (for example, half the residents saw an anterior MI with a 55-year-old banker, and the other half with an 80-year-old woman; at test all saw a 55-year-old banker with a left bundle branch block. Accuracy of residents dropped from 46% to 23% when the test case historical data matched the prior (incorrect) example. The fact that the effect was observed with manipulation of information that was not diagnostically relevant suggests that the process was not available to critical reflection, meeting 1 characteristic of exemplar memory.

These studies suggest that, in at least some domains, reasoning proceeds by early identification of possible diagnoses through recognition of a similar prior example: the 'hypothesis generation' phase of the early descriptive studies. This cannot be the only way to generate hypotheses, as for rare conditions prior examples may have arisen in the distant past or not at all. Nevertheless, the studies do point to the existence of an experience-based knowledge that is very different from the analytical representations discussed earlier.

HOW CAN WE TEACH CLINICAL REASONING – OR CAN WE?

What are the implications of these research findings for teaching clinical reasoning? Although this research agenda began with the objective of revealing a reasoning process used by experts so that it could be taught to students, no reasoning process that accumulated with expertise emerged. Instead, all the research we have reviewed suggests that expert clinical reasoning is a consequence of an extensive and multidimensional knowledge base. A direct consequence is that any attempt to teach a process of clinical reasoning that does not emphasise the centrality of knowledge, is, in Eva's[36] words:

> 'outdated and inaccurate… Reasoning ability is not a ''trait'' that can be assigned to an individual… the context within which a problem is being addressed has a major impact on the accuracy of the decisions reached.'

Does that mean that there are no shortcuts, and becoming an expert simply amounts to acquiring all that knowledge? Or is it the case that some strategies of knowledge use are more effective than others? A few studies have examined the relative effectiveness of various strategies and knowledge representations in reaching a correct outcome; with various methods, both observational and experimental. In examining these studies, it is worthwhile to keep in mind the earlier caveat that showing that something is associated with success does not imply that it causes success.

Probably the earliest example of differing strategies derived from the observational studies of Patel and Groen,[7] who claimed to show that successful problem-solvers and/or experts used a process of 'forward reasoning' from data to diagnosis (e.g. the patient has crushing retrosternal chest pain with radiation down the left arm. It is a likely myocardial infarction), whereas novices and less successful problem-solvers use backward reasoning ('It could be an MI because it is crushing in character. On the other hand, it might be endocarditis since she had an infection last week'). However, Eva *et al.*[37] showed that this was probably an epiphenomenon, confounded with the use of *post hoc* probes and the observational nature of the data; it is not clear that success leads to forward reasoning on subsequent explanation of the case or forward reasoning leads to success. Further, an experimental study[38] where students were encouraged to use either forward or backward reasoning in diagnosing ECGs actually showed a marked superiority for a combined approach, beginning with backward reasoning.

Elieson and Papa[39] contrasted learning probabilities of symptoms in various diseases with learning lists of features with qualifiers such as 'usually'. They showed a difference in favour of probability learning, with scores of about 70%, vs. 58% for those who received only the verbal probability descriptors. However, probabilities are not easy to remember, and Woods[21] used the same materials, but with 1 group learning probabilities while a second group learned basic science explanations of the features. They showed that, while there was no difference on immediate test, after 1 week a group that had learned basic science mechanisms outperformed the probability group by about 10%.

In several observational studies, use of SQs has also been associated with successful problem-solving, leading to an experimental study of an instructional intervention.[40] However, while the experimental group showed substantially increased use of SQs, diagnostic accuracy was about the same in both groups.

Coderre *et al.*[27] investigated the extent to which novice and expert problem-solvers used schema induction vs. 2 other strategies: pattern recognition and hypothetico-deductive reasoning. Schema induction was associated with 5 times greater success and pattern recognition with a 10 times advantage, over hypothesis generation. Surprisingly, the authors conclude that:

> [Pattern recognition's] use by medical students is not usually advocated because their inadequate experience might lead to potentially grim consequences. (p. 699)

Because the study is observational these results, like Patel and Groen's,[7] may confound the direction of causality. Although the authors are aware of the problem and attempt to argue that this is unlikely, it cannot be ruled out.

On the other hand, it is plausible that some domains are simply too complex to permit simple pattern recognition strategies. A recent study by McLaughlin *et al.*[41] in nephrology showed that schema induction was a common and effective strategy for both experts and novices (although experts also used non-analytical processes effectively), which is consistent with another study in nephrology.[20]

Finally, Kulatanga-Moruzi *et al.*[42] trained medical students in dermatological diagnosis, then at test encouraged 1 group to use pattern recognition and the other to carefully identify features, then arrive at a diagnosis. There was no difference in accuracy between the 2 groups. A more recent study used a third group that began with pattern recognition, then checked with systematic search strategy. This group had about a 10% increase in accuracy.

There is a paradox that the field began with a search for general strategies, then refocused on the type and organisation of knowledge. But the general strategies never quite went away; they simple transmuted. Instead of 'problem-solving' and 'reasoning', they became 'pattern recognition', 'schema induction' or, 'Bayesian inference'. But it must be recognised that these are not general, content-independent, strategies; they are, instead, strategies to access and apply different kinds of specific knowledge. As a consequence, instruction on how to use a strategy in the absence of a parallel focus on the kind of knowledge needed to apply the strategy is unlikely to succeed, as the studies of instructional interventions above highlight. Perhaps more important than practice with a specific strategy is deliberate practice itself, an educational strategy whose critical contribution to mastery and application of concepts has not been adequately recognised.[14,43,44] When one changes the focus from testing one or another reasoning strategy within a problem to a concern with the optimal sequencing of problems to maximise learning, preliminary research has shown that large gains in performance are possible.[45,46]

Taken together, these findings lead to 2 conclusions. First, it is possible to observe different strategies in knowledge use, which often lead to markedly different outcomes. However, when subjected to experimental manipulation, differences are small or non-significant. Whether this reflects our previous concerns that clinicians, both novice and expert, have access to and use different forms of knowledge concurrently, or whether it simply indicates that these processes are not easily amenable to instruction, is not clear. However, the more important finding is that focusing instruction on 1 processing strategy or another may be less important than engaging students with many problems, which are carefully sequenced to optimise learning and transfer.

HOW CAN WE ASSESS CLINICAL REASONING 'SKILLS'?

A direct consequence of the review is that there is no single, or even optimal pathway through a problem.

We cannot, for example, anticipate that a successful solution implies that the student induced a schema, generated a correct hypothesis or worked through a Bayesian probability matrix. Indeed, with increasing expertise, success requires less mental effort. As Anderson[47] said:

> 'One becomes an expert by making routine what to the novice requires creative problem-solving ability.' (p. 292)

The implications for assessment are clear. Given the poor correlation across problems observed in multiple studies[48] and the many potential pathways to a correct solution, a preoccupation with assessment of the reasoning process, requiring intense scrutiny of one or a few problems[49] is misguided. Instead, simply determining successful outcome on multiple short problems is much more defensible.

CONCLUSIONS

Research on clinical reasoning began with the goal of understanding the process of reasoning. It became clear early on that the process was almost too universal, in that everyone, young or old, novice or expert, was doing approximately the same thing. The field then moved away from a process focus to a critical examination of expert knowledge. While this shift in direction has clearly led to new insights, the search for *the* knowledge organisation has proved elusive. While it is appealing to presume that an expert has some kind of superior organisation of knowledge and/or superior strategies or skills, it is now clear that an expert possesses superior knowledge of many kinds, both formal and informal, and any or all may be brought to bear on the solution of a particular problem.

While this appears to add complexity to the task of teaching and assessing clinical reasoning, it can be seen to result in both simplification and reorientation. It is not of particular consequence that a student is provided with specific kinds of knowledge or specific strategies (e.g. forward reasoning, use of semantic qualifiers, pattern recognition). Rather, it is emerging that central to the acquisition of expertise, both in medicine and many other domains[1,4,44] is the opportunity for deliberate practice with multiple examples and feedback, both to facilitate effective transfer of basic concepts and to ensure an adequate experiential knowledge base.

Is the acquisition of expertise simply a matter of sufficient practice? If so, this would represent a dramatic about-turn from where the field began, with a concern with the identification of general strategies used by experts. Further, if that was all to it then there is little left to research, and the acquisition of expertise resembles the response of a New Yorker to a tourist's question, 'How do I get to Carnegie Hall?'.

The answer: 'Practice, practice, practice'.

However, that is not all there is. It has become clear that experts use multiple knowledge representations in solving a problem, and the kind of knowledge brought to bear is more critical of success than the process. For straightforward and frequently encountered problems, similarity-based reasoning is undoubtedly effective and efficient. At the other end, when problems are rare and complex, the expert is able to marshal an extensive array of scientific and experiential knowledge;[20] but one clearly unanswered question is how the expert switches from one to another; how does the expert recognise that the problem does not fit the mould? A facile response might be that this is where reflective practice[50] comes in, but as a recently published article pointed out, 'no empirical research has been conducted to date into the nature of reflective practice in medicine'.[50,51]

Finally, we end where we begin. What has been called 'expert clinical reasoning' really amounts to expert diagnostic reasoning, usually in internal medicine. We know little about the factors that influence management decisions. Advocates of evidence-based medicine and decision analysis methods can tell us much about what *should* influence management decisions (at least, according to their particular world view), but can tell us little about how practitioners actually weigh up the many factors, medical, social and psychological to arrive at a particular course of action.

We also know little about other domains of medicine. There is a growing literature related to surgical expertise, which has been excluded from this discussion[52] and which introduces many other factors – manual dexterity, visual–spatial co-ordination, into discussions of expertise. Radiology has its own, albeit limited, literature[53] and reasoning in areas such as anaesthesiology and critical care medicine is very different again resembling, in part, the vigilance of the aircraft pilot and in part the fine 'tweaking' of a complex non-linear system that one sees in an expert mechanic.

One thing is clear. There is no such thing as clinical reasoning; there is no one best way through a problem. The more one studies the clinical expert, the more one marvels at the complex and multidimensional components of knowledge and skill that she or he brings to bear on the problem, and the amazing adaptability she must possess to achieve the goal of effective care.

Contributor: Geoffrey Norman.
Acknowledgements: the author wishes to acknowledge the contribution of the Natural Science and Engineering Research Council and the Canada Research Chair Program in supporting this research.
Funding: part of this work was supported by the Government of Canada under the Canada Research Chair program and a grant from the Natural Science and Engineering Research Program.
Ethical approval: none.
Conflicts of interest: none

REFERENCES

1 McGuire CH. Medical problem-solving: a critique of the literature. *J Med Educ* 1985;**60**:587–95.
2 Elstein AS, Shulman LS, Sprafka SA. *Medical Problem Solving: an Analysis of Clinical Reasoning*. Cambridge, MA: Harvard University Press 1978.
3 Neufeld VR, Norman GR, Barrows HS, Feightner JW. Clinical problem-solving by medical students: a longitudinal and cross-sectional analysis. *Med Educ* 1981;**15**:315–22.
4 Barrows HS, Norman GR, Neufeld VR, Feightner JW. The clinical reasoning process of randomly selected physicians in general medical practice. *Clin Invest Med* 1982;**5**:49–56.
5 Simon HA, Chase WG. Skill at chess. *Am Scient* 1973, 1983;**61**:394–403.
6 Muzzin LJ, Norman GR, Feightner JW, Tugwell P. Expertise in recall of clinical protocols in two specialty areas. Proceedings of the 22nd Conference on Research in Medical Education, Washington, DC: 1983;122–6.
7 Patel VL, Groen GJ. Knowledge based solution strategies in medical reasoning. *Cogn Sci* 1986;**10**:91–115.
8 Schmidt HG, Boshuizen HPA. Transitory stages in the development of medical expertise: the 'intermediate effect' in clinical case presentations. *Proceedings of the Cognitive Science Society*. Hillsdale, NJ:. Erlbaum ;139–145.
9 Norman GR, Brooks LR, Allen SW. Recall by experts and novices as a record of processing attention. *J Exp Psychol: Learn, Mem Cognit* 1989;**15**:1166–74.
10 Schmidt HG, Boshuisen HPA. On the origin of intermediate effects in clinical case recall. *Mem Cogn* 1993;**21**:338–51.
11 Verkoeijen PPJL, Rikers RMJP, Schmidt HG, van der Weil MWJ, Koomna JP. Case presentation by medical experts, intermediates and novices for laboratory data presented with or without a clinical context. *Med Educ* 2004;**38**:617–27.
12 Burns BD. The effects of speed on skilled chess performance. *Psychol Sci* 2004;**15**:442–7.
13 Gruppen LD, Frohna AZ. Clinical reasoning. In: Norman GR, Newble DI, Van der Vleuten CPM, eds. *International Handbook of Research in Medical Education*. Dordrecht: Kluwer 2002.
14 Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Acad Med* 2004, 1988;**79**: S1–12.
15 Day SC, Norcini JJ, Webster GD, Viner ED, Chirico AM. The effect of changes in medical knowledge on examination performance at the time of recertification. Proceedings of the 27th Conference in Research into Medical Education, Chicago, IL: 1988;139–42.
16 Schmidt HG, Norman GR, Boshuizen HPA. A cognitive perspective on medical expertise:. theory implications. *Acad Med* 1990;**65**:611–21.
17 Boshuizen HPA, Schmidt HG. Biomedical knowledge and clinical expertise. *Cogn Sci* 1992;**16**:153–84.
18 Patel VL, Evans DA, Groen GJ. Biomedical knowledge and clinical reasoning. In: Evans DA, Patel VL, eds. *Cognitive Science in Medicine*. Cambridge, MA: MIT Press 1988.
19 Patel VL, Groen GJ, Norman GR. Effects of conventional and problem based medical curricula on problem solving. *Acad Med* 1991;**66**:380–9.
20 Norman GR, Trott AL, Brooks LR, Smith EKM. Cognitive differences in clinical reasoning related to postgraduate training. *Teach Learn Med* 1994;**6**: 114–20.
21 Woods NN, Brooks LR, Norman GR. The value of basic science. *Med Educ* 2005;**39**:107–12.
22 Shank RC, Abelson RP. *Scripts, Plans, Goals and Understanding: An Inquiry Inot Human Knowledge Structures.* Hillsdale NJ, Erlbaum 1977.
23 Barrows HS, Feltovich PJ. The clinical reasoning process. *Med Educ* 1987;**21**:86–91.
24 Charlin B, Roy L, Brailovsky C, Goulet F, van der Vleuten C. The Script Concordance test: a tool to assess the reflective clinician. *Teach Learn Med* 2000;**12**:189–95.
25 Papa FJ, Shores JH, Meyer S. Effects of pattern matching, pattern discrimination, and experience in the development of diagnostic expertise. *Acad Med* 1990;**65**:S21–2.
26 Mandin H, Jones A, Woloschuk W, Harasym P. Helping students learn to think like experts when solving clinical problems. *Acad Med* 1997;**72**:173–9.
27 Coderre S, Mandin H, Harasym PH, Fick GH. Diagnostic reasoning strategies and diagnostic success. *Med Educ* 2003;**37**:695–703.
28 Mandin H, Harasym P, Eagle C, Watanabe M. Developing a 'clinical presentation' curriculum at the University of Calgary. *Acad Med* 1995;**70**:186–93.

29 Bordage G, Lemieux M. Semantic structures and diagnostic thinking of experts and novices. *Acad Med* 1991;**66**:S70–2.
30 Bordage G, Connell KJ, Chang RW, Gecht MR, Sinacore JM. Assessing the semantic content of clinical case presentations: studies of reliability and concurrent validity. *Acad Med* 1997;**72**:S37–9.
31 Brooks LR. Decentralized control of categorization: the role of prior processing episodes. In: Neisser U, ed. *Concepts and Conceptual Development.* Cambridge: Cambridge University Press 1978;141–74.
32 Medin DL, Altom MW, Murphy TD. Given versus induced category representations: use of prototype and exemplar information in classification. *J Exp Psychol Learn Mem Cogn* 1984;**10**:333–52.
33 Regehr G, Cline J, Norman GR, Brooks LR. Effect of processing strategy on diagnostic skill in dermatology. *Acad Med* 1994;**69**:S34–S36.
34 Kulatanga- Moruzi C, Brooks LR, Norman GR. Coordination of analytical and similarity based processing strategies and expertise in dermatological diagnosis. *Teach Learn Med* 2001;**13**:110–6.
35 Hatala RM, Norman GR, Brooks LR. Influence of a single example upon subsequent electrocardiogram interpretation. *Teach Learn Med* 1999;**11**:110–7.
36 Eva KW. What every teacher needs to know about clinical reasoning. *Med Educ* 2005;**39**:98–106.
37 Eva KW, Norman GR, Brooks LR. Forward reasoning as a hallmark of expertise in medicine: logical, psychological, phenomenological inconsistencies. *Adv Psychol Res* 2002;**8**:25–40.
38 Norman GR, Brooks LR, Colle CL, Hatala RM. The benefit of diagnostic hypotheses in clinical reasoning: experimental study of an instructional intervention for forward and backward reasoning. *Cogn Instruct* 2000;**17**:433–48.
39 Elieson SW, Papa FJ. The effects of various knowledge formats on diagnostic performance. *Acad Med* 1994;**69**:S81–S83.
40 Nendaz MR, Bordage G. Promoting diagnostic problem representation. *Med Educ* 2002;**36**:760–6.
41 McLaughlin K, Coderre S, Mortis G, Mandin H. Variables affecting the choice of diagnostic reasoning strategies and diagnostic success in novices and experts. 12th Ottawa Conference, July 2004, Barcelona.
42 Kulatunga-Moruzi C, Brooks LR, Norman GR. The diagnostic disadvantage of having all the facts: using comprehensive feature lists to bias medical diagnosis. *J Exp Psychol Learn Mem Cogn* 2004, **30**, 563–72.
43 Eva KW, Neville AJ, Norman GR. Exploring the etiology of content specificity: factors influencing analogical transfer and problem solving. *Acad Med* 1998;**73**:S1–S6.
44 Guest CB, Regehr G, Tiberius RG. The life long challenge of expertise. *Med Educ* 2001;**35**:78–81.
45 Hatala RM, Norman GR, Brooks LR. Practice makes perfect. the critical role of deliberate practice in the acquisition of ECG interpretation skills. *Adv Health Sci Educ* 2003;**8**:17–26.
46 Avrahami J, Kareev Y, Bogot Y, Caspi R *et al.* Teaching by examples: implications for the process of category acquisition. *Q J Exp Psychol* 1997;**50A**:586–606.
47 Anderson JR. *Cognitive Psychology and its Implications.* San Francisco, CA: Freeman 1980.
48 Eva KW. On the generality of specificity. *Med Educ* 2003, 1979;**37**:587–8.
49 Painvin C, Norman GR, Neufeld VR, Walker I, Whelan G. The triple-jump exercise – a structured measure of problem-solving and self-directed learning. Proceedings of the 18th Conference on Research in Medical Education, Washington, DC; 1979:73–7.
50 Schon DA. *The Reflective Practitioner: How Professionals Think in Action.* New York: Basic Books 1983.
51 Mamede S, Schmidt HG. The structure of reflective practice in medicine. *Med Educ* 2004;**38**:1302–6.
52 Matsumoto ED, Hamstra SJ, Radomski SB, Cusimano MD. The effect of bench model fidelity on endourologic skills: a randomized controlled study. *J Urol* 2002;**167**:1243–7.
53 Norman GR. Expertise in visual diagnosis: a review of the literature. *Acad Med* 1992;**67**:S78–S84.

Received 4 October 2004; editorial comments to author 8 December 2004; accepted for publication 18 January 2005