Expert/Novice differences in Diagnostic Medical Cognition - A Review of the Literature

L. Cuthbert [1], B. duBoulay[1], D. Teather[2], B. Teather[2], M. Sharples[3] & G. duBoulay[4]

[1] School of Cognitive and Computing Sciences, University of Sussex
[2] Department of Medical Statistics, De Montfort University, Leicester
[3] School of Electronic and Electrical Engineering, University of Birmingham
[4] Institute of Neurology, London

University of Sussex

Cognitive Science Research Paper

CSRP 508

February 1999

ISSN 1350-3162

Abstract: *The nature of expertise has been studied in a wide variety of disciplines. Many of these studies have found that the development of expertise involves qualitative as well as quantitative changes in the cognitive skills and knowledge representations underlying performance. This paper seeks to review the literature which has sought to identify expert/novice differences in the field of diagnostic medical cognition. This literature is concerned with the way experts and novices make diagnostic medical decisions. Differences are identified in terms of hypothesis generation and testing, diagnostic reasoning and the organisation of relevant knowledge. Furthermore, expert's diagnostic reasoning is characterised as largely schema driven, with previous patient encounters influencing medical evaluations of current cases.*

Medical cognition refers to studies of cognitive processes (such as perception, comprehension, decision making and problem solving) in medical practice. Studies have tended to concentrate on the decision making processes and have tended to ignore other areas of medical expertise (such as patient management and reporting), and it should be acknowledged that this gives a rather narrow view of medical expertise. Studies of expertise in diagnostic medical cognition examine differences between practitioners with different levels of experience in terms of their cognitive processes and skills. Many of the studies (e.g. Patel, Arocha & Kaufman, 1994) distinguish between *novices* (individuals who have only everyday knowledge of a domain or the pre-requisite knowledge assumed by the domain, i.e. medical students), *intermediates* (individuals who are above the beginner level but below the sub-expert level, for example, medical residents), *sub-experts* (individuals with generic knowledge but inadequate specialised knowledge of the domain, for example, cardiology experts solving problems in the area of endocrinology) and *experts* (an individual with specialised knowledge of the domain, for example, cardiology experts solving cardiology problems). Some studies have broken these categories down even further, for example, by distinguishing between novices with different levels of experience (*early, intermediate and advanced novices* referring to 1st, 2nd and 3rd year students respectively - Arocha & Patel, 1995), or between *basic-experts* and *super-experts* (Raufaste, Eyrolle & Marine, 1998). Super-experts refer to the top experts in a particular field.

The aim of this paper is to provide a review of the literature examining expertise effects on diagnostic medical cognition. Generally speaking, different studies have concentrated on different aspects of expertise effects on cognition, including hypothesis generation and evaluation, memory performance, diagnostic reasoning and the organisation of clinical knowledge. These issues will be dealt with independently in this paper. The paper also examines the more perceptually skilled discipline of radiology. Diagnosis in radiology involves a two step process. First, the radiologists must perform a visual search of the radiograph to examine whether any abnormalities are present in the image. Second, if any abnormalities are identified by the scan, then the radiologist must further examine the areas of abnormality in order to perform a diagnosis. Hence, radiology can be seen as a perceptually skilled discipline. If abnormalities are missed in the first scan of the image, then diagnosis may not be complete and accurate. We may therefore expect radiologists to possess different cognitive skills than experts specialising in domains of medicine with a lesser visual component. This paper does not however concentrate on the visual perceptual skills per se of radiologists - instead the emphasis is on the second stage, the diagnostic process itself, and on the interaction between perception and problem solving.

Most studies addressing the issue of expertise in medical cognition have tended to use the same basic experimental paradigm in which subjects are shown a written description of a clinical case (Patel & Groen, 1986, Patel, Groen & Arocha, 1990) and instructed to read the case notes for a specific period of time. The notes are then removed and the subject asked to recall details of the case and then to provide a diagnosis. In a variation of this paradigm, subjects are presented with the clinical information sequentially, one sentence at a time and asked to explain the incoming information and suggest a diagnosis. These methods are known as the *immediate presentation paradigm* and the *sequential presentation paradigm* respectively. The studies then use semantic, propositional and conceptual analysis techniques to characterise the structure of the subject's knowledge, to examine differences in problem representation and finally to evaluate levels of coherence in the subject's response (see Patel, Arocha & Kaufman (1994) for further detail). Think aloud instructions are also often included. Some researchers have questioned the extent to which these rather artificial experimental tasks allow any valid conclusions to be drawn with regard to the nature of the diagnostic process. Patel, Evans & Kaufman (1989) moved away from the written clinical presentations and instead examined doctor/patient interviews. Medical practitioners were asked to interview patients with the goal of developing a differential diagnosis. Patients were paid volunteer outpatients who had recently presented with the history and symptoms of an endocrine disorder (the study required the diagnosis of an endocrine disorder). The interviews were conducted in actual settings with realistic time constraints. Physicians were allowed to ask the patient any questions they liked in producing their diagnosis. Although this kind of study can be seen as having more ecological validity than the laboratory based studies discussed earlier, there is still an argument that even this type of study fails to capture the real processes undertaken in producing a diagnosis. Indeed, some researchers (e.g. Klein, Calderwood & MacGregor, 1989, Huber, 1997) believe that in examining real world decisions a more naturalistic approach needs to be taken. Such an approach is based on the belief that real world decisions are influenced by the context in which they are made, studying decisions outside their real world context is not a valid approach. Despite their limitations, laboratory studies of diagnostic medical cognition have produced some interesting results. The aim of this paper is to summarise and discuss these major findings.

Expertise effects

One of the earliest stages in the diagnostic process involves the formulation of working hypotheses. These are derived from clinical observations (in the form of signs, symptoms, test results, and results of physical examinations). The first section of the literature review examines the issue of the effects of expertise on hypothesis generation and evaluation.

1. Hypothesis Generation & Evaluation.

Sisson, Donnelly, Hess and Woolliscroft (1991) examined how experts (with various different areas of sub-speciality within Internal Medicine) and novices (3rd year medical students) differ in terms of the number, specificity and breadth (range) of diagnostic hypotheses generated early in the evaluative process. They argue that these initial hypotheses, although modified by subsequent data, will have a large influence on the diagnostic process. Their results showed that on average, the medical students generated more hypotheses than did the physicians, and this finding was consistent across scenarios. Furthermore, physicians' hypotheses were found to be more general than the students; however, the two groups did not differ in terms of the breadth of their hypotheses (calculated as a percentage of possible categories of diagnosis). Neither group typically named all of the diagnostic categories that logically might have been included as part of their hypotheses. Sisson et al. assume that the fact that physicians generate less specific hypotheses (in fact, their hypotheses were quite general) may reflect a consciously learned approach or an intuitive evolution in reasoning. The problem with highly specific diagnoses based on limited data is that they may result in some form of premature closure, hence specific hypotheses can be seen as a pitfall that reduces the options available to lead to the proper diagnosis. A final, perhaps surprising finding was that the numbers of hypotheses generated by the individual participants was fairly consistent across the three tasks studied. Other research suggests that the speed with which initial hypotheses are generated is a striking feature of the behaviour of experts. Furthermore, there is evidence that the earlier a good hypothesis set is created, the more predictive it is of the quality of the diagnosis (Joseph & Patel, 1990) .

Lesgold, Rubinson, Feltovich, Glaser, Klopfer & Wang (1988) found that novice (resident) radiologists examining chest x-rays appear to restrict their responses to the most obvious explanation. Two hypotheses were put forward to explain these findings: first, it may be that when there is a dominant hypothesis and a more remote possibility; consideration of the more remote possibility depends upon the availability of mental processing capacity. If any subprocesses of diagnosis are inefficient they will interfere with the more remote response. An alternative explanation is that novices simply do not generate the full range of sensible possibilities in forms that will survive testing and verification. Novices may have learned the triggering rule for the most obvious explanation but not for the subtle special cases. Novices may fail because they have not yet developed the fine-tuned visual acuity needed for feature discrimination that is seen in their more experienced colleagues.

Patel, Arocha & Kaufman (1994), Joseph & Patel (1990) found that experts (endocrinologists solving endocrine problems) produce their hypotheses fairly quickly and accommodate subsequently presented data without introducing any new hypotheses. In contrast, sub-experts (cardiologists solving endocrine problems) continue to generate new hypotheses even after producing most of the diagnostic components needed for the final diagnosis. They are less able to evaluate their hypotheses and hence show an inability to rule out diagnostic hypotheses they had produced earlier. In general, the experts narrowed uncertainty whereas the sub-experts increased it. They suggest that sub-experts do not have sufficient domain knowledge to discriminate hypotheses. In contrast, experts' initial hypothesis set appears to be particularly well constructed allowing them to complete the diagnostic task from the hypotheses contained in the original set without adding new hypotheses.

Arocha & Patel (1995) examined the effects of inconsistent data on subjects' hypotheses generation and evaluation. Early, intermediate and advanced novices (2nd, 3rd & 4th year medical students) showed differences in terms of their use of co-ordinating operations. These are responses to inconsistent data that are commonly used in scientific and everyday reasoning and include; ignoring data, excluding data, re-interpreting data, re-interpreting hypothesis, modifying a hypothesis to fit the data, and changing a hypothesis altogether. Early and intermediate novices performed more data operations and they more frequently ignored or reinterpreted inconsistent data, whereas advanced novices more often changed their hypotheses to account for the data, changes which decreased the inconsistency with the data. Advanced novices generated a number of early hypotheses and this allowed them to narrow their initial hypothesis set in the face of inconsistent evidence and make fewer data reinterpretations. Intermediates also generated a number of early hypotheses, however they failed to change their hypotheses when confronted with inconsistent data. Instead they maintained several hypotheses of a diverse nature concurrently without evaluating them efficiently. Arocha & Patel concluded that training appears to make heir hypothesie ofng aace nees wh work (through their research interests and diagnosing unusual cases) would trigger sufficient levels of deliberate reasoning - the kind necessary to constrain possibilities optimally in such cases. Such a hypothesis is consistent with Ericsson, Krampe & Tesch-Romer's (1993) view that 'eminent performance' is directly related to the 'amount of deliberate practice related to that goal'. Interestingly, the authors also found that the richness of the semantic networks underlying problem solving was better able to explain diagnostic success than years of experience.

Another important area of research has examined the effects of expertise on the directionality of reasoning employed. This refers to the distinction between forward reasoning (in which the physician attempts to generate a hypothesis from the findings in a case) and backward reasoning (in which the data collection is influenced by the hypothesis generated). Patel & Groen (1986) found that directionality of reasoning is related to diagnostic accuracy. They found that the diagnostic explanations of subjects making an accurate diagnoses (of acute bacterial endocarditis) consisted of pure forward reasoning. In contrast, subjects with inaccurate diagnoses tended to make use of forward and backward reasoning. Subsequent studies have confirmed this relationship between diagnostic accuracy and forward reasoning. Patel, Groen & Arocha (1990) examined factors which may disrupt this pattern of forward reasoning. They examined the reasoning strategies used by expert cardiologists and endocrinologists solving problems within or outside their area of expertise. They found that an accurate diagnosis (usually from the experts) was associated with pure forward reasoning in the production of an explanation of the principle component of the diagnosis. However, this was often accompanied by one or two components of backward reasoning to explain any loose ends. In contrast, inaccurate diagnoses (usually from the sub-experts) were associated with the nd back.rndwits wf- lly fromaruptioni (loose ends. In dang0 2gnostic explaeen fpothwo componee exoductdng. inacom to knowledgts)Howevererts tsno built-to checksuosltgiitiec a(a gforoune baosknowled

links to relate critical or relevant cues, showing better organisation of their domain knowledge. Similar findings have been obtained in radiology. Lesgold et al. (1988) found that experts reported more different findings, had longer reasoning chains, larger and more clusters, and a greater number of their findings connected to at least one other finding. Hence experts differ from novices in terms of the coherence of their knowledge and explanations.

As expected, on most tasks experts show superior diagnostic performance in classifying clinical conditions. Furthermore, there is evidence to suggest that the superior diagnostic skills of experts may be due to their ability to fully utilise patient information presented in clinical protocols (known as enabling conditions). Schmidt, Hobus, Patel & Boshuizen (1987) found that when expert family doctors were presented with slides containing a picture of the patient (allowing them to deduce the subjects age, sex etc), and information about the patients profession, previous diseases, medication, marital status and so forth, then they showed superior diagnostic skills (compared to novice family doctors) when they were subsequently presented with the patient’s complaint (38% vs. 27%). They suggest that this may be due to differential use of enabling conditions information as the experts remembered 40% more of this information. Hobus, Hofstra, Boshuizen & Schmidt (1989) found that when information about enabling conditions was not presented to expert physicians, their diagnostic performance was no better than that of novice physicians. They also found that when experts were asked to describe typical patients with certain diseases, they gave richer descriptions, in particular with reference to the contextual factors facilitating the emergence of the disease.

Similar findings have been obtained with experts in radiology. Norman, Brooks, Coblentz & Babcook (1992) examined the effects of information contained in a brief clinical history on diagnosis and on feature identification from radiographs. They were concerned with the diagnosis of bronchilitis, and a bronchilitis history (fever, cough and tachypnea) was induced for some trials. Having been presented with the case, paediatric radiologists were asked to rate the likelihood of bronchilitis on a scale from -3 for definitely absent to +3 for definitely present. Feature history was found to affect both diagnosis (with a change in the scale of between one half and one unit on the scale), and feature identification (the effect amounted to an increase of about 25%-50% in the number of features identified on film). Furthermore, no novice-expert differences were found in terms of the ratings of likelihood of

(forceful feature) which gave rise to these interpretations. They failed to find any differences between experts (registrars and consultants in general medicine) and novices (1st & 3rd year medical students) in terms of the number of interpretations provided, and this finding lead them to conclude that level of clinical experience does not affect the number of interpretations made in response to any given array of clinical information. However, they did find differences in the content of thought. A large amount of variability was found not just between groups but also within them. In every case there was found to be a small focal area of common interpretation and a massive peripheral field of individual differences. No differences in the number of forceful features were identified, but differences in the actual forceful features themselves were found. There was found to be little overlap, both between and within groups. When they examined the forceful features associated with the correct diagnosis, they found that the more difficult diagnoses allowed identification of far fewer forceful features than the easy diagnoses. They concluded that individuals develop various ways of accessing their memory structures in the case of easy diagnoses, however people do not demonstrate such extensive networks for difficult diagnoses. Highly individualised multiple responses to an array of clinical information are associated with ease of diagnosis. When a diagnosis is being missed, few structures to the right memory structure are found amongst those who think of it and none amongst those who don't. They concluded that experience is not characterised by uniformity of thinking but by individuality of thought. In diagnostic thinking there is no best way, no key piece of information. Experts vary greatly as do students, the difference is that experts have used and changed the knowledge stored so that it becomes progressively more useful personally.

Moskowitz, Kuipers & Kassirer (1988) suggest that the use of heuristics dominates experts clinical problem solving and probabilistic judgments. They distinguish between categorical descriptions of likelihood (e.g. low, moderate, high), ordinal descriptions of likelihood (where quantities are described in terms of ordinal relations, such as greater or less than a reference value) and numerical descriptions of likelihood (where quantities are described in terms of numerical measures). They found that experts references to likelihood consisted primarily of the first two (categorical and ordinal - where numerical values were used, they were used to define categories (70-80%) or as an anchoring value for an ordinal description). Furthermore, the widespread use of two heuristics (representativeness and anchoring) was found. Tied to categorical descriptions of likelihood are archetypic examples,

distinguishing features. They speculate that the process of diagnosis involves the physician focusing on relatively small sets of logically related diseases, (i.e. small worlds) and carrying out a limited number of comparisons among these diseases. It is hypothesised that diseases contained within these 'small worlds' would typically share certain overlapping features, and this is the basis for their membership of that particular 'small world'. However, the diseases contained within a 'small world' differ in terms of the presence or absence of certain other features, allowing the expert to distinguish between the candidate diseases contained within a 'small world'. In order to investigate the 'small worlds' hypothesis, Kushniruk et al. re-analysed the protocols collected in the studies carried out by Joseph & Patel (1990) and Patel, Evans & Kaufman (1989). They were particularly concerned with examining the networks of relationships among the hypotheses and findings generated by experts (endocrinologists solving endocrine cases) and sub-experts (cardiologists solving endocrine cases). The networks produced by the experts were found to contain few elements (i.e. a limited number of hypotheses and findings) which were tightly connected, displaying a high degree of coherence and relatedness. Furthermore, expert physicians quickly focused on those cues and critical findings in a medical case that most clearly distinguish among competing diagnoses in the hypothesis set under consideration. In contrast, the sets of hypotheses generated by sub-experts often contain large numbers of diagnostic hypotheses each belonging to different disease categories. Kushniruk et al. argue that an expert's knowledge is organised in this way because of the limitations of human memory and processing capacity. Furthermore, this organisation of knowledge affects the way experts perform a diagnosis, (i.e. the experts' knowledge organisation and reasoning processes are viewed as being integrally related).

The 'small worlds' hypothesis is closely related to the concept of schema, indeed many theories of expertise assume that experts processing of clinical cases is schema driven. Schemas are defined in cognitive psychology as hypothetical cognitive structures that allow us to call upon our past experience and knowledge in interpreting the present situation. Schemas are thought to play an important role in facilitating the recognition of significant objects within a problem and in enhancing the ability to recognize typical situations. The presence of schemas is used to explain an expert clinician's ability to pay attention to relevant information only and to diagnose diseases rapidly. According to the 'small worlds' hypothesis physicians manage large amounts of information by restructuring their knowledge into small sets of logically related disease schemas. This allows experts to use more efficient discriminatory strategies in order to rule out competing hypotheses, allowing them to focus on the few acquiring expertise, students progress through several transitory stages, characterised by distinctively different knowledge structures underlying their performance. Second, they assume that these representations do not decay or become inert in the course of developing expertise but remain available for future use, when the situation requires activation. Finally, the model assumes that experienced physicians, whilst diagnosing routine cases, use knowledge structures or ’illness scripts’. These structures develop through continuous exposure to patients. They contain little knowledge of pathophysiological causes of symptoms and complaints - but a wealth of clinically relevant information about the disease, its consequences and the context in which it develops. There are various levels of illness scripts from categories of diseases to individual patients seen before.

According to this ’stage’ model, early in the development of medical knowledge, students develop rich, elaborated causal networks explaining the causes and consequences of disease in terms of general underlying pathophysiological processes. Early learning is based largely on books, and this results in a prototypical perspective on disease, with limited understanding of variability of the disease in reality. At the second stage, brought about through extensive and repeated exposure to real patients, the declarative knowledge outlined in stage 1 becomes compiled into high-level, simplified causal models explaining signs and symptoms, which are subsumed under diagnostic labels. Short cuts become available through experience. Students at this level of development will not have to activate all relevant knowledge in order to understand the patient, only knowledge pertinent to the case. In short, knowledge is re-organised to ensure accessibility and efficient use, into simplified causal models explaining signs and symptoms.

The third stage is reached once the student has had sufficient experience meeting ’real’ patients. As a result of meeting patients, the student gets a feeling of how disease manifestations may vary, paying attention to contextual factors under which disease emerges. Instead of causal factors, the different features that characterise the clinical appearance of a disease becomes the anchor point around which their thinking evolves. The student develops a series of representations called ’Illness Scripts’. These structures contain varied information including: enabling conditions - factors making occurrence of a disease more likely (e.g. hereditary factors), including predisposing factors (e.g. drugs) and boundary conditions (e.g. age, sex)., faults - a description of the malfunction, and consequences - the signs and symptoms arising from the fault. Furthermore, illness scripts provide the rules enabling one to construct mental models of a family of diseases (based on feature overlap - in this way the model has parallels with the small worlds hypothesis), a specific disease or even a patient having the disease. Illness scripts also seem to specify a specific order of information - they obey certain conventions regarding an optimal structure in medicine. Coughlin & Patel (1987) found that experts' (family practitioners) memory for case information was more sensitive than novices (2nd year medical students) to the ordering of the data presented. They suggest that experience results in experts expecting information to be presented in a certain order, if the information presented is not consistent with such expectations, then this affects their ability to process the information. Illness scripts, because they develop through experience are highly idiosyncratic in nature. For any disease, an individual physicians scripts may or may not resemble the scripts of other physicians or the textbook. The absence of pathophysiological information in the scripts (apart from a simple model explaining the causes) implies this information is not normally used by the physician. Indeed as discussed earlier, there is evidence that experts do not use this form of data in their diagnostic reasoning.

Finally, having reached the final stage, experts begin to store patient encounters as 'Instance Scripts'. According to this hypothesis, memory for previous patient encounters are retained in memory as individual entities and are not merged into some prototypical form. Furthermore, an assumption is made that, to a large extent, expert clinical reasoning is based on the similarity between the presenting situation and some previous patient available from memory. Hassebrock & Pretula (1990) found that physicians retained vivid autobiographical memory for cases seen as long as 20 years earlier. Van Rossum & Bender (1990) showed the impact of a single vivid case on diagnosis of similar examples two years later. Pattern recognition plays an important part in diagnosis.

Central to the model is the assumption that as clinicians gain experience and develop new information processing structures, previously acquired knowledge remains available and expert physicians may move from one stage to another as the complexity of the problem demands. The model does not suggest that experts work at a deeper level of processing, but rather, that expertise is associated with the availability of knowledge representations in various forms, derived from both experience and formal education.

Schmidt & Boshuizen (1993) describe an elaborated model of expertise development

immediately, and cause currently active plans to be interrupted or abandoned in favour of new exploratory activity. Rogers noted at least two different types of attentional activity. The first is characterised by relatively fast noticing and labelling of an abnormality as soon as the x-ray

The evidence suggests that memory measures are a less reliable measure of expertise. Specifically, unlike diagnostic reasoning, recall is nonmonotonically related to expertise. As a general rule, intermediates perform better on memory tasks than either experts or novices, and it is suggested that this is because they process more of the case material that is not relevant to the diagnosis. This processing results in superior memory for a larger number of case details, however because this information is not relevant to the diagnosis, it is not related to diagnostic success. Experts and sub-experts cannot be distinguished in terms of memory measures.

In terms of diagnostic reasoning, experts have been found to make extensive use of information contained within the case notes (known as enabling conditions), furthermore this finding has been found to extend to both general family practitioners (Schmidt, Hobus, Patel & Boshuizen, 1987) and to the more specialised discipline of radiology (Norman, Brooks, Coblentz & Babcook, 1992). However, experts make less use of biomedical terms and explanations in their diagnosis, instead their explanations are based on clinically relevant information. Clinical relevant information is specified by experts' own representation of the disease known as an 'illness script'. 'Forceful features' refer to those personally important pieces of information which give rise to a clinical interpretation. Little consistency has been found in terms of the forceful features associated with a diagnosis by experts or novices. Hence, it seems that experience is not characterised by uniformity of thinking but by individuality of thought, experts vary greatly as do students, the difference is that experts, through experience, have used and changed the knowledge stored so that is becomes progressively more useful personally.

Experts appear to organise their clinical knowledge of different disease in terms of small groups of confusable diseases (known as 'small worlds'). These groups have in common certain important features which is the basis for their membership of any particular small world. These structures also contain information which allows the clinician to distinguish between the competing diagnostic hypotheses contained within that small world, thus aiding correct diagnosis. Experts have also been found to retain vivid memories of previous patient encounters, in the form of 'instance scripts'. Previous cases diagnosed by an expert have been found to affect diagnosis of a similar case upto 20 years later, with clinical reasoning based on the similarity between the two cases.

Despite the fact that studies of diagnostic medical cognition have shown clear differences in the cognitive skills and knowledge organisation possessed by experts and novices, Johnson (1988) points out that research which has focused upon performance presents a rather pessimistic appraisal of experts, with experts often failing to perform any better than novices. For example, Goldberg (1959) found no differences between psychiatrists and their secretaries in terms of their ability to diagnose brain damage using a common test. Furthermore, Goldberg (1968) found that a regression model performed with greater accuracy than experts in the diagnosis of psychosis using the MMPI (Minnesota Multiphasic Personality Inventory). Johnson (1988) examined physicians' ability to evaluate applicants for internships. He found that a simple regression model was more effective in selecting successful candidates than the experts. To explain these findings, he point out that experts often include in their analysis individual case data that would not be included in a regression model, in other words, experts seem to pay attention to relatively rare variables, applicable only to the case under consideration (these are termed 'broken leg cues'). HowefeaHcexperts thesincormanion ' (oued by eegression model s). HThs idsticncion isanalyogousto the cdsticncion etween p we discussed earlier how experts, when judging likelihood, make widespread use of two heuristics, the representativeness heuristic and the anchoring heuristic (Moskowitz et al, 1988). Furthermore, Johnson’s (1988) characterisation of experts as being particularly influenced by case-specific data is consistent with the findings of Schmidt et al. (1988), who characterised experts as being particularly influenced by the patient information contained in the clinical protocols (the enabling conditions). Hence, there is further evidence to support Johnson’s claims.

Studies like Johnson’s which examine error rates of experts and statistical models have however, been criticised as they fail to take account of the type and consequences of different types of error. Hence it may be that whilst statistical models make fewer errors than experts, the errors the models do make may prove to be more serious (i.e. there may be graver consequences for the patient). This would then prove to be a mitigating factor against the use of statistical models in diagnosis.

A related issue affecting practitioners’ ability to produce an accurate diagnosis concerns their use of base rate probabilities. In particular, we can distinguish between two probabilities relevant to diagnosis, firstly, the probability of disease from base rates, and secondly, the probability of disease from the clinical presentation of the case. Accurate diagnosis will be dependent upon a physicians ability to incorporate case data accurately and modify diagnostic probabilities accordingly. There is also a wider issue concerning whether practitioners should use base rates at all - or simply consider the probability of disease from the presentation of the case. Perhaps the strongest argument that can be made in favour of ignoring base rates is that they penalise rare complaints in favour of more common illnesses.

To conclude, studies of diagnostic medical cognition suggest that experts differ from novices and sub-experts in terms of a number of skills which affect their diagnostic accuracy. In particular, experts are proposed to have better or more complete representations of the task domain, better organisation of task relevant knowledge and superior diagnostic reasoning skills. However, it seems that such a characterisation of experts may not tell the whole story. It seems that experts’ reasoning is characterised by the widespread use of various heuristics and influenced by biased data evaluations. The finding that experts tend to underweigh base rate information compared to case-specific information should be approached with caution however, as the task employed by Johnson (1988) was not strictly speaking a medical one.

However, the conclusion is given some credibility by the studies discussed earlier in this report which suggest that expert physicians' diagnostic reasoning is strongly influenced by the enabling conditions presented in the clinical data. As pointed out by Johnson (1988), the implications of these findings are of particular relevance to practitioners involved in the design of decision support tools. It is clear that these systems should examine the information that experts currently underweigh and provide a measure of its impact. The expert can then adjust this initial estimate to account for information not considered by the model, such as the enabling conditions, hence taking advantage of the computational abilities of the computer and the skills and experience of expert physicians.

References

Adelson, B. (1984). When novices surpass experts: The difficulty of a task may increase with expertise. Journal of Experimental Psychology, 10, 483-495.

Arocha, J.F., & Patel, V.L. (1995). Novice diagnostic reasoning in medicine: Accounting for Evidence. The Journal of the Learning Sciences, 4(4), 355-384.

Arocha, J.F., Patel, V.L., & Patel, Y.C. (1993) Hypothesis generation and the coordination of theory and evidence in novice diagnostic reasoning. Medical Decision Making, 13, 198-213.

Boshuizen, H.P.A., Schmidt, H.G., & Coughlin, L.D. (1988). On the application of medical basic science knowledge in clinical reasoning: Implications for structural knowledge differences between experts and novices. Proceedings of the 10th Conference of the Cognitive Science Society, 517-523. Hillsdale, New Jersey: Erlbaum.

Boshuizen, H.P.A. & Schmidt, H.G. (1992). The role of biomedical knowledge in clinical reasoning by experts, intermediates and novices. Cognitive Science, 16, 153-184

Brooks, L.R., Norman, G.R & Allen, S.W. (1991). Role of specific similarity in a medical diagnostic task. Journal of Experimental Psychology: General, 120, 278-287.

Charness, N. (1979). Components of skill in bridge. Canadian Journal of Psychology, 33, 1-16.

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

Cochran, A.L. (1972). Effectiveness and Efficiency. Random Reflections on Health Services. London: Nuffield Provincial Hospitals Trust. (Reprinted in 1989 in association with the British Medical Journal).

Coughlin, L., & Patel, V.L. (1987). Processing of critical information by physicians and

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

Hassebrock, F. & Pretula, M. (1990). Autobiographical Memory in medical problem solving. Paper presented at the American Educational Research Association Meeting, Boston, Massachusetts.

Hillard, A., Myles-Worsley M., Johnson, W. & Baxter, B. (1985). The development of Radiologic schemata through training and experience: A preliminary communication. Investigative Radiology, 18 (4), 422-425.

Hobus, P.P.M., Hofstra, M.L., Boshuizen, H.P.A., & Schmidt H.G. (1989). Mental representation of prototypical patients: Expert-Novice differences. Paper presented at the First European Congress on Psychology, Amsterdam.

Huber, O., Wider, R, & Huber, O.W. (1997). Active information search and complete information presentation in naturalistic risky decision tasks. Acta Psychologica, 95(1), 15-29.

Johnson, E.J. (1988). Expertise and decision under uncertainty: Performance and process. In Chi, M., Glaser, R., & Farr, M.J. (Eds). The Nature of Expertise, 209-228, Hillsdale, NJ: Erlbaum.

Joseph, G.M., & Patel, V.L. (1990). Domain knowledge and hypothesis generation in diagnostic reasoning. Medical Decision Making, 10, 31-46.

Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. Cognitive Science, 12, 1-55.

Klein, G., Calderwood, R., & MacGregor, D. (1989). Critical decision method for eliciting knowledge, IEEE Trans. Syst., Man Cybern. 19(3), 462-472.

Kushniruk, A.W., Patel, V.L. & Marley, A.J. (1998). Small worlds and medical expertise: implications for medical cognition and knowledge engineering. International Journal of Medical Informatics, 49, 255-271.

Lesgold, A. (1984). Acquiring expertise. In J.R. Anderson & S.M. Kosslyn (Eds.), Tutorials in learning and memory: Essays in honour of Gordon Bower (pp.31-60). San Francisco: W.H. Freeman.

Muzzin, L.J., Norman, G.R., Jacoby, L.L., Freightner, J.W., Tugwell, P., & Guyatt. G.H. (1982). Manifestations of expertise in recall of clinical protocols. In Proceedings of the 21st

Patel, V.L., Evans, D., & Kaufman, D. (1989). A cognitive framework for doctor-patient interaction, in: D. Evans, V. Patel (Eds.), Cognitive Science in Medicine: Biomedical Modelling, MIT Press, Cambridge, MA, pp. 253-307.

Patel, V.L., Groen, G.J., & Arocha, J.F. (1990). Medical expertise as a function of task difficulty. Memory & Cognition, 18(4), 394-406.

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

Raufaste, E., Eyrolle, H & Marine, C. (1998). Pertinence Generation in Radiological Diagnosis: Spreading Activation and the Nature of Expertise, Cognitive Science, 22(4), 517-546.

Regehr, G. & Norman, G.R. (1996). Issues in Cognitive Psychology: Implications for Professional Education. Academic Medicine. 71(9), 998-1001.

Rogers, E. (1995). A Cognitive Theory of Visual Interaction. In B. Chandrasekaran, J. Glasgow & N. Hari Narayanan (Eds.), Diagrammatic Reasoning: Cognitive and Computational Perspectives, AAAI Press/ The MIT Press, Menlo Park, California.

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

Schmidt, H., Norman, G., & Boshuizen, H.P.A. (1990). A cognitive perspective on medical expertise: Theory and implications. Academic Medicine, 65(10), 611-621.

Schmidt, H.G., & Boshuizen, H.P.A. (1993). On the origin of the intermediate effect in clinical case recall. Memory & Cognition, 21, 338-351.

Schmidt, H.G., & Boshuizen, H.P.A. (1993). On Acquiring Expertise in Medicine, Educational Psychology Review, 5, 205-221.

Sisson, J.C., Donnelly, M.B., Hess, G.E., & Woolliscroft, J.O. (1991). The characterisation of early diagnostic hypotheses generated by physicians (experts) and students (novices) at one medical school. Academic Medicine, 66, 607-612.

Van Rossum, H.J.M., & Bender. W.W. (1990). What can be learnt from a boy with Acute Appendicitis? Persistent effects of a case presentation on the diagnostic judgement of family doctors. Paper presented at the Fourth Ottawa Conference, Ottawa, Ontario, Canada.

Weber, E.U., Bockenholt, U., Hilton, D.J., & Wallace, B. (1993). Determinants of diagnostic hypothesis generation: Effects of information, base rates and experience. Journal of Experimental Psychology: Learning, Memory & Cognition, 19(5), 1151-1164.