Developing the Attributes of Medical Professional Judgement and Competence

Michael Eraut
Institute of Education
University of Sussex

Benedict du Boulay
School of Cognitive and Computing Sciences
University of Sussex

April 7, 2000

This literature review was funded by the Postregistration Medical and Dental Education Research Initiative of the Department of Health's Policy Research Programme. This version is available on the web at `http://www.cogs.susx.ac.uk/users/bend/doh`

Contents

Executive Summary

Introduction

1. Descriptive data about the processes of British postgraduate medical education is sparse. Outcomes evidence is limited to pass rates for the Royal Colleges examinations and the individual judgements of assessment panels for GP registrars and completion of CCST.

2. Postgraduate trainees are employed as working health professionals, so their current competence is always an issue, as well as the competence they will be expected to demonstrate when their specialist training has been completed. Defining competence in terms of the expectations of the holder of a particular post helps to avoid confusion.

3. The central problem of postgraduate education is how best to combine work within a doctor's current competence, itself a source of learning, with the provision and use of learning opportunities to extend that competence.

4. There is a need for an agreed framework for the development of competence during each training programme, which puts learning outcomes on the priority list alongside service duties.

5. Areas in which competence is seriously underconceptualised include communications, teamwork and management in healthcare settings.

6. The recent report of the US Federated Council for Internal Medicine Task Force, Graduate Education in Internal Medicine: a Resource Guide to Curriculum Development ([SEKR97]) is a useful source of ideas.

7. Some perceive judgement as an attribute of personal expertise that goes beyond that competence which any fully trained doctor could be reliably expected to demonstrate. It can also be seen as a dimension of lifelong learning linked mainly to the improvement of decision-making through learning from experience over a long period, rather than the learning of new practices or keeping up to date with research.

Research into Medical Expertise and Decision-Making

8. Key features of expertise include the importance of case-based experience, the rapid retrieval of information from memory attributable to its superior organisation, the development of standard patterns of reasoning and problem-solving, quick recognition of which approach to use and when, awareness of bias and fallibility; and the ability to track down, evaluate and use evidence from research and case-specific data. Understanding the nature of expertise is important for self-monitoring one's use of heuristics and possible bias, sharing knowledge with others and supporting other people's learning. It is also critical for understanding the respective roles of clinical experience and evidence-based guidelines.

9. Research into decision-making under conditions of stress and uncertainty suggests that training in crisis management is needed, and that teamwork and other organisational factors are important.

10. There is a need for regular self-evaluation to maintain critical control of one's practice.

11. The use of evidence-based medicine requires on-the-job as well as off-the-job teaching.

Learning in Clinical Settings

12. Many features of the educational policy seem to be appropriate, but they are not being implemented in many hospitals. There is insufficient supervision and feedback. Educational goals are subordinated to service demands. While many house officers receive good clinical teaching, a minority do not and assurance of educational quality is weak. Learning goals are only specified at a very general level, so there is little clarity about priorities, especially at the PRHO stage.

13. This issue has to be tackled at local level where there is limited management of the educational process and clinical tutors have little time and no authority over clinical teaching. Deans do what they can; but quality assurance of postgraduate education lags well behind that for clinical practice; and the UK research base at this level is minuscule.

14. The learning of procedures in medical posts has been criticised for being too haphazard: there is often little continuity of experience and guidance is often provided by doctors who are themselves not very experienced.

15. The appropriateness for GPs of so much general hospital training has been questioned. Though research on this issue would be difficult, we think more research evidence could and should be gathered.

16. The key issues emerging from North American research are:

    - the time allocated for trainer-trainee interaction and making the best use of that time;

- methods of providing feedback;
- qualities of teachers rated as 'good' by trainees and their colleagues.

17. One gets the impression that American residents receive considerably more clinical teaching than their British counterparts, but there is no British data to enable a proper comparison. The variation in the amount of training received by British trainees is reported as considerable, raising issues of quality assurance and trainee entitlement.

Learning in Non-Clinical Settings

18. Research into postgraduate teaching and learning in non-clinical settings mostly comprises evaluations of a wide variety of teaching innovations, rich in ideas but not necessarily generalisable.

19. Strong evidence that the effectiveness of off-the-job teaching is highly dependent on its links with related on-the-job teaching makes it unwise to evaluate off-the-job teaching on its own.

Continuing Medical Education and Lifelong Learning

20. Surveys of GPs, and also in a few cases consultants, have shown the importance for learning and changes in practice of a wide range of learning activities and sources of information. Moreover, they differ according to whether the changes involve treatment (including prescription), diagnosis and investigation, doctor-patient relationships, referral policy, health promotion or practice organisation. Models of physician learning distinguish between learning triggered by the problems raised by current individual patients and "learning projects" to acquire or improve proficiency in a targeted area of practice. The initiation of learning is dependent on significant background knowledge of what is out there to be learned to which CME conversations with other physicians, and reading contribute in ways which would not be revealed, for example, by evaluations of CME events. The importance of informal consultations with others and a reluctance to "cold call" experts suggest that facilitating social interaction among doctors and strengthening their networks should be a policy goal.

21. Evaluations of CME courses have demonstrated the importance of including activities such as the observation and discussion of visual material and/or supervised practical work. Though it has confirmed that short courses of 1 day or less are rarely effective, no controlled studies have been reported which used length of course as a variable. This deficiency needs to be remedied because much time could be wasted trying to improve courses which are too short; and unrealistic expectations of the learning time required for certain goals are easily developed by busy learners and under-resourced providers — a form of collusion from which nobody benefits. Another important conclusion is that educational interventions on their own often fail to achieve changes in practice.

22. Research on innovation strategies points to the danger of focusing only on the development of competence. Competence has to be translated into performance and at this stage many dispositional and organisational factors come into play. Research on the implementation of guidelines, for example, indicates not only that the quality and utility of the guidelines themselves is important but also that both educational interventions (leading to understanding of their purpose and rationale) and administrative interventions (ranging from organisational changes to simple reminders) need to accompany the guidelines.

23. The discussion of recent developments in CPD reaches two conclusions. First, needs analysis is important for quality assurance purposes at three levels - the individual, the working group and the healthcare organisation (the last two are multi-professional). However, it should not be assumed that needs identified by audit, for example, will necessarily require an educational response. Second, following the advice of [FB98b], CME providers should adopt a coordinated approach to all three levels by facilitating self-directed learning, providing high quality individual and group education, and assisting healthcare organisations to develop and practise organisational learning.

The Role of Information Technology

24. Decision support systems have had a mixed reception over the years, but seem to be growing in acceptance as just another tool in the doctor's armoury. A training issue here is the need to ensure that doctors understand how decision support systems frame the problem so that they can judge the quality of the advice that they offer.

25. In terms of the development of such systems, we may expect the development of linked databases rather than fragmented sources and better explanations from systems about the reasons for their decision advice.

26. The training implication of this is that such systems may be excellent, but their use needs to be carefully integrated into the overall training programme — not least so that the human trainers provide what the computer-based system cannot provide in terms of monitoring and feedback.

Assessment and Feedback

27. Summative assessment for the award of the Certificate of Completion of Specialist Training will be at least partly performance-based; and will need to be evaluated as it comes on line. Since Membership Examinations of the Royal Colleges are competence-based, the performance-based element at the end of Basic training is given relatively little attention. Thus the most critical certification issue arising from our review is the extent to which assessment regimes cover the full range of competence discussed in Chapter 3 and its translation into performance.

28. With formative assessment questions have been raised about frequency, coverage, and reliability; and, if it is to properly serve its purpose, the manner in which formative assessment is integrated into training programmes to support the learning process will also need to be researched. Levels of supervision are often affected by factors other than the competence level of the trainee, and feedback may not be based on any systematic (though not necessarily formal) assessment.

29. Learning to assess and reassess one's own competence and its limits is a long and complex process, which becomes increasingly sophisticated as a doctor progresses through postgraduate education. Its reliability significantly depends on access to good supervision and feedback. Feedback which contributes to a trainee doctor's self-assessment may come from patient outcomes, informal discussions with other doctors or health professionals, periodic appraisals or meetings for signing the trainee's logbook. Informal feedback on the wards tends to be spontaneous and incidental, i.e. not the result of a reflective judgement: it is mostly received from more senior trainees rather than consultants, and more likely to be negative than positive. Formal feedback appears to vary considerably in quantity, quality and breadth of coverage. Even the best designed log-books focus on competence rather than performance.

Monitoring Postgraduate Education

30. As reported in Chapter 4, there is sufficient cause for concern to suggest that research into the practice and effectiveness of supervision and feedback during postgraduate medical education is urgently needed. This should include the implicit delegation of certain supervisory and educational responsibilities to senior trainees. Should it be formalised, as in the US role of Chief Resident? Should it be accompanied by training? Which responsibilities could or should be delegated, and which should not?

31. Few Trusts have reliable mechanisms for incorporating clinical teaching into their organisation of professional time, though some may have formally agreed to do so. Nor are there any internal audit mechanisms at local level for monitoring and periodically evaluating a Trust's programme of professional education. Clinical Tutors have neither the time nor the authority to undertake such duties.

Mentoring after Qualification

32. Given the difficulties discussed above of finding sufficient time for supervision, giving feedback and clinical teaching for postgraduate medical trainees, the introduction of mentoring as an additional role and obligation might not justify a high priority. However, mentoring might be particularly well suited to the support of doctors during the first few years after completion of postgraduate training. Both consultants and GPs could benefit from such support as they grow into their new roles and responsibilities, learn to work with new colleagues and to contribute appropriately to their Trust or General Practice, and take greater responsibility for organising their own lifelong learning.

Chapter 1

Introduction

1.1 Aim of the Review

The increasing codification and regulation of medical competence highlights the multi-faceted and complex nature of being a doctor. Not only are extensive skills and knowledge expected within the doctor's area of specialism, but also high levels of communicative ability, ethical understanding and responsibility, teamworking capability and organizational ability.

Although much is learned pre-registration, the greater part of a doctor's overall and effective competence and judgement is developed post-registration: through active involvement in a large number of cases; through teaching and supervision; through discussion and teamwork; through reflection on practice; and through formal study.

Two questions underpin this review. First, what is the nature of medical competence and judgement? Second, how can post-registration medical education be organized so as to develop it in an effective and efficient manner? However the first question is answered, it is clear that the pivotal learning experience is, and will remain, the doctor's exposure to, and involvement in, a wide range of real cases; cases where difficult judgements have to be made, often with uncertain and insufficient data and sometimes against severe time pressures.

1.2 Methods Used

The main method used was a review of the research literature in two fields:

1. professional competence, judgement and expertise in medicine; and
2. post graduate and continuing medical education.

Where appropriate, research literature from outside the field of medicine has been introduced which relates to professional expertise, the definition and specification of competence, decision-making, and learning in the workplace: these are fields with which either or both of the authors is conversant. Details of the search methodology for (1)

and (2) above are provided in Appendix A. The bulk of this search was conducted by Ian Eiloart, the researcher on the project, whose assistance we gratefully acknowledge.

In addition to the search, we interviewed 12 experienced doctors with responsibilities in Postgraduate Medical Education, and studied Royal Colleges' documents on Basic Training, Higher Specialist Training and the Vocational Training of General Practitioners. The main purposes of the interviews were:

- to ascertain our respondents' understanding of the terms **competence** and **judgement**;
- to find out what they considered to be critical factors affecting the success of postgraduate medical education, and the policy issues which had been addressed by the recent reforms or still needed to be addressed.

This helped us to contextualise our report, and confirmed that knowledge of the research under review was not widely distributed.

Given the substantial discussion of major papers in this report and the provision of abstracts for most entries in our database, we decided that the provision of an additional annotated bibliography would be redundant. The database itself will be available through the website of the University of Sussex School of Cognitive and Computing Sciences, as will the final (after review) text of this report.

1.3 Reporting of Research

Methodological approaches to research in medical education include a rather different balance of methods than those found in biomedical research. While this difference can be partly attributed to research funding on a much smaller scale, there are also many practical constraints on research into medical education. Numbers at postgraduate level are often too small for the use of inferential statistics, so a combination of qualitative methods and descriptive statistics is often used. Allocation to experimental or control groups is normally both random and explicit, but cannot often be blind. Since doctors in training are fairly sophisticated subjects, qualitative methods can allow some attribution of cause to desirable or undesirable outcomes, in the absence of such control studies.

Whenever control studies have been used we have noted this fact in the text, but we have not tried to make any judgements about whether they have been well used. Before and after studies are also reported as such. The absence of such annotations can be reliably interpreted as signifying the absence of such methods. There have been a small number of systematic reviews, mostly focussed on continuing medical education, where [DTOH95]'s review is the most frequently cited. [Irb95]'s review on learning in ambulatory settings is another milestone, which includes a wider range of research methods reflecting the constraints and research feasibility in that area. On the theoretical side, research into clinical expertise tends to be based on written case descriptions and protocols, with decisions emphasising at the possible expense of natural validity.

1.4 The structure of this Report

Following this Introduction are two chapters on professional competence, judgement and expertise in medicine. Chapter 2 provides a definition of **competence** based on the expectations of employers, patients and the medical profession. This approach is implicit in the General Medical Council's publication **Good Medical Practice**; and is particularly well suited to postgraduate training, during which doctors' service contributions need to match their current competence, in addition to providing learning opportunities to extend that competence. This is followed by a brief discussion of the term **judgement** which we have not attempted to define. There are at least three meaning in use: some psychologists treat the term as synonymous with decision-making; our informants used it either with reference to its quasi-legal role in assessment or to refer to particularly complex decisions involving a high level of uncertainty. Instead of attempting to resolve these differences, we have focused on the term **expertise**, which has been the subject of considerable research. Chapter 3 reviews research into medical expertise, which seems to be almost universally defined in terms of the capability of those people deemed by their peers to be experts, or at least highly experienced and well respected. Thus the term "expertise" subsumes both competence and clinical judgement. Our understanding is that Postgraduate Medical Education is mandated to develop the competence expected of consultants of GP principals; and expected to make a significant start to the development of expertise which will continue to develop through lifelong learning long after postgraduate training has been completed.

Chapters 4 and 5 review the research on learning and the support of learning through guidance, coaching and supervision during postgraduate medical education. Chapter 4 covers learning in clinical settings and Chapter 5 learning in non-clinical settings. The majority of this research has been conducted in North America, where postgraduate medical education is based on **residency programmes**: these last for 3 to 6 years after graduation from Medical School, according to the specialism. Give the increasing emphasis on the maintenance of competence and further development of expertise after the completion of training, we have added a further chapter on Continuing Medical Education (CME) and Continuing Professional Development (CPD). Chapter 6 is particularly important because of the large volume of recent research on how doctors learn and the factors affecting changes in their practice; and preparation for this lifelong learning phase is an important goal for postgraduate training programmes.

Chapter 7 is a short specialist chapter on the use of Information Technology in training doctors, an increasingly important source of innovation in medical education. Chapter 8 reviews a second specialist area, Assessment and Revalidation, because these processes set the minimum standards for postgraduate medical education and continuing medical education, respectively. The "backwash effect" of assessment on learning is one of the most firmly established findings of educational research; so this has to be given constant attention as well as the normal expectations of valid and reliable assessment practices.

Finally, Chapter 9 discusses the implications of this research review for policy: the proper use of the construct of competence and the efficient and effective provision of post-registration medical education.

To aid the reader some synopses of key papers are "boxed" in the text. These boxed sections are a continuous part of the main text, though they can be read on their own.

Chapter 2

Competence and Judgement

2.1 Competence

The term competence appears in many forms and guises in the research literature on the professions. Nevertheless, a clear definition is needed to guide a review of this kind. [Era98]'s review of definitions and meanings of competence distinguishes between those authors who treat competence as a socially situated concept — the ability to perform tasks and roles to the standard people expect — and those who define it as individually situated, a set of personal capabilities or characteristics ([NN85]). Eraut argues in favour of a socially situated definition, because the notion of competence is central to the relationship between professionals and their clients; and recommends using the word **capability** to describe the individually situated concept of "what a person can think or do". Whether or not a person's capability makes them competent in a particular job depends on them being able to meet the requirements of that job. Hence **competence** in a job is defined as **the ability to perform the tasks and roles required to the expected standard**.

The advantage of this definition is that it can be applied to a professional at any stage in their career, not only to the newly qualified. The standard expected will no doubt vary with experience and responsibility and take into account the need to keep up to date with changes in practice. It also leaves open the question of who will decide what is to count as competence when different people have different expectations. This is an essentially political issue. While the GMC has the legal authority, many cases are 'resolved' at local level or between doctor and patient without referring the matter any further. The position of postgraduate trainees is more complicated as they are still under supervision; nevertheless it would clearly be inappropriate for those towards the end of their training to be closely supervised all the time. Training results in a steadily increasing range of competence, accompanied by gradually decreasing levels of supervision; and the process of expanding one's range of competence continues after completion of training. Throughout this period the principle holds good of not undertaking work for which one is not competent without appropriate supervision, whatever one's status.

The GMC's publication **Good Medical Practice** states in a preface entitled "Duties and Responsibilities of Doctors" that:

> "The principles of good medical practice and the standards of compe-

> tence, care and conduct expected of you in all aspects of your professional work are described in this booklet. They apply to all doctors involved in healthcare." (page 1)

This confirms the importance of meeting expectations and emphasises that this applies to **all aspects of professional work**, not just those that might be described as purely medical. The GMC does not refer only to competence but also to **care** and **conduct**, because there can be a gap between **competence** — what a doctor can do - and **performance** — what a doctor does do. This gap can be caused by personal factors such as dispositions or attitudes, environmental factors such as workload, workplace climate and working conditions, or situational factors such as multiple emergencies and lack of support. While membership examinations for the Royal Colleges assess medical knowledge and clinical competence, assessments during specialist training and for CCST are based on performance on the job over a considerable period of time.

[RvLD...90] point out that many researchers have failed to appreciate this distinction between competence and performance, also noting that only one of the eight studies reviewed was able to show that competence was a significant predictor of performance. They argue that this was not due to measurement deficiencies, as claimed, because it was a behavioural problem rather than a psychometric problem. More research should focus on the conditions which influence doctors' performance. Their own study ([RSD...91]) of 36 General Practitioners in the Netherlands used standardised patients both for an initial assessment of competence under relaxed conditions and for blind visits to their practices 12 months later to assess performance under normal working conditions. When considering assessment results alone, the doctors did better on the competence than the performance assessment. When using efficiency-time scores, the doctors did better in actual practice. Only when both assessment results and efficiency-time scores were used did competence predict performance. In a different context, [RBF...91]'s analysis of 45 anaesthesiology residents handling of critical incidents showed that two thirds of the variability was predicted by two non-cognitive variables, "conscientiousness" and "composure". Thus below standard performance may not signify lack of competence.

2.2 Specifying Competence

Precisely what constitutes the **expected standard** which defines **competence** can also be a problem. The most basic specification of competence is that used by the General Medical Council, whose new assessment arrangements have now been tested in the courts. Then for General Practitioners there are two further specifications to be considered. The Summative Assessment of GP Registrars is set at a higher standard than the GMC basic; and at a higher standard still are the requirements for becoming a Member of the Royal College of General Practitioners. This is also intended to be achieved towards, or soon after, the end of the training period. The confusion this causes the public can be traced to two separate uses of the term competence in everyday discourse. One usage treats the attribution of competence to a person as a binary decision: either one is competent or one is not competent. The decision might not be easy but only two answers are possible. For example, for any given situation one is either insured or uninsured. The other usage treats competence as a position on a continuum of expertise, probably somewhere in the middle, i.e. rather less than excellent. One

might be very content with a competent lawyer when buying a house, but look for an expert when seeking advice on a piece of complex litigation. Thus, being referred to as "competent" could in some circumstances be perceived as a rather negative comment.

[Pie93b] explores the practical implications of multiple levels of competence, both confirming the need for a minimum level of performance, below which behaviour is defined as negligent (i.e. the GMC level for keeping on the register) and arguing for targeting postgraduate education and training on a level well above that minimum. He also points out that setting the standard below the current average would be complacent, setting it at the average would be conservative and setting it above the average would be very challenging. Whatever the standard of performance at the end of training, the expectation must be for continuing improvement thereafter; and this depends on the doctor's attitude towards quality improvement and lifelong learning. Many of those we consulted commented on the gap between the standard associated with the Certificate of Completion of Specialist Training (CCST) and that expected of an established consultant, and this must be at least as significant for GPs who receive much less training for a job of similar complexity. Thus the chapter on Continuing Medical Education is an important part of this report.

Specifications of competence are essential for setting standards, and the Certificate of Completion of Specialist Training (CCST) is competence-based. So also are the clinical components of membership examinations of the Royal Colleges. Nevertheless, the transition from implicit judgements of a doctor's competence to explicit descriptions which can guide both training and assessment has not been easy. Many would argue that it still has a long way to go. Difficulties include:

1. finding the most appropriate level of detail for defining a policy, an educational programme, a summative assessment, or a log-book;
2. capturing the essence of an area of expertise and not being content with the most easily observable aspects;
3. both listing important attributes of competence and describing their integration into performance;
4. covering all aspects of a doctor's job, not only the diagnosis and treatment of individual patients.

These problems are not confined to the field of medicine. Nevertheless, the importance for educational policies and programmes, which aim to develop the attributes of competence and judgement, of specifying both the attributes and the nature of competence itself cannot be underestimated. Apart from the publications of the Royal Colleges themselves, there are a number of journal articles discussing various aspects of competence and a range of publications from the Royal College counterparts in North America and Australia.

One document which we found particularly valuable, both in its description of aspects of competence and its discussion of the design of appropriate postgraduate programmes to develop competence is the recent report of the Federated Council for Internal Medicine Task Force on the Internal Medicine Residency Curriculum ([SEKR97])

which is available from the American College of Physicians and also on the Internet `www.asim.org/fcim`.

In order to overcome some of the difficulties listed above, this Task Force asked experts in the field of Internal Medicine "to identify those difficult-to-measure properties of the expert internist that might be overlooked after biomedical knowledge and skills are presented in organ-specific competency lists." They called this construct "physicianship" and saw it as "the difference between the expert internist and the journeyman practitioner". They grouped the suggestions received into 20 categories, which they call Integrative Disciplines; and these are accompanied by 24 categories of clinical competency.

The term "discipline" may not seem quite appropriate in the British context, but the categories themselves make good sense. These are listed in the Panel below, according to the three FCIM tiers of category. First, are the **core values of internal medicine** which, they anticipate, will be learned largely through discussion, problem-solving exercises, workshops and practical experience. Next, are the **characteristics, or salient expressions**, of these values, which are likely to require at least some formal instruction in addition to opportunities for practice and formal feedback. The final tier comprises **integrative skills, or applications**, of the values and characteristics, for which theory may be less important and the main approach to learning is acquiring experience in realistic settings.

Integrative Disciplines and Categories Comprising Physicianship

1. Core Values Humanism Professionalism Medical Ethics	**2. Characteristics/Expressions of Core Values** Lifelong Learning Clinical Method Continuity of Care Medical Interview/History Physical Diagnosis Clinical Epidemiology and Quantitative Clinical Reasoning Clinical Pharmacology Scientific Literacy Legal Medicine Management of Quality of Health Care Nutrition Preventive Medicine
3. Integrative Applications Home care Nursing Home Care Occupational/Environmental Medicine Physical Medicine/Rehabilitation Management of Medical Practice Medical Information	

[SEKR97]

For each of these headings 6-12 competencies are listed and possible learning venues suggested. The examples listed below indicate the general approach. In essence this document takes a set of duties similar to those listed in the GMC's Good Medical Practice and provides a set of competencies to indicate what a doctor would need to understand and do in order to fulfil those duties with a high level of professionalism.

Examples of General Competencies Associated with Physicianship

---

Humanistic Practice of Medicine (page 25)
Recognise and appropriately manage so-called "difficult patients," including their personality disorders and problematic behavioural patterns.
Understand the concepts of the health belief model; know how to elicit it and how to work constructively in a patient-centred way with persons from different cultural groups.
Professionalism (page 27)
Show a commitment to standards for lifelong excellence by continuously adding to one's knowledge of medicine and by drawing the distinction between knowledge that is based on high-quality evidence and knowledge from anecdote and personal experience.
Medical Ethics (page 29)
Know what to do when the patient refuses a recommended medical intervention in both emergency and non-emergency situations. Know how to handle the following situations related to end-of-life care:

- Withholding or withdrawing life-sustaining treatment, including nutrition and hydration
- Communicating "bad news" and listening for the patient's and family's concerns
- Writing "do not resuscitate" orders
- Requests for physician-assisted suicide or euthanasia
- Know how to address requests to breach confidentiality
- Know the principle of truth-telling and how to implement it in situations involving information disclosure and medical errors

The Medical Interview (page 39)
Use the interview to identify cognitive impairment, anxiety, denial, and defensiveness; be able to manage each during the interview.

[SEKR97]

Another approach to mapping the range of competence is to focus on areas of activity and responsibility, and relationships with the other people involved.

For example, six areas of responsibility can be usefully distinguished:

1. The acquisition and interpretation of evidence about the patient's health, concerns and living environment;
2. The care and management of the patient and of the whole health care episode;
3. Communication with the patient, the care group, and other relevant parties;
4. Monitoring and evaluation of the patient's health, the care and treatment given,

and the administration and management of the service;

5. Quality improvement: quality of life for the patient, quality of practice for health care workers and teams, quality of service for the health care organisation; and

6. Health promotion and preventive medicine.

The main groups with whom doctors have to develop and sustain appropriate working relationships are

- The patients whom they see
- Their family and friends
- The care group of health workers who have contact with those patients
- Other doctors in the health care organization or relevant parts of it
- the health care organization itself

These responsibilities and relationships determine the expectations of doctors and hence the range of their competence. Then individual doctors also have overall responsibility for controlling their own behaviour, for self-management and self-evaluation. We do not propose to describe in detail the full range of clinical skill required for each specialism, though some are much more clearly described than others. That would go beyond our capability and involve too much specialist discussion for a report whose prime focus is learning. Instead we shall elaborate further the range of competence associated with what [CTN...99] recently described as "the next and potentially more challenging stage" in the development of Higher Specialist Training:

> "Ensuring that, in addition to developing effective clinical skills, trainees are also able to respond to service changes and to develop a wider range of competencies, including team working, communication skills, and the ability to identify health needs and understand the opportunities for health promotion." (page 30)

2.3 Communicative Competence

Often the least well defined area in the doctor-patient interaction zone is that of communication. The GMC booklet **Good Medical Practice** sets out the doctor's responsibilities very clearly in paragraph 12. A general disposition to be polite, considerate and respectful is important; but there are also quite complex skills to be learned. Patient rights to know and to choose are of central concern; but the general principle of **informed consent** begs the questions of 'how well informed' and 'whose consent'. [MWMT92] investigated the influence of treatment descriptions on elderly patients completing 'advance directives'. 77% changed their minds at least once when given the same scenario but a different description of the intervention, of whom two thirds changed their mind four or more times. The FCIM Task Force suggests that a doctor should:

- "Be able to assess a patient's decision-making capacity;
- Know how to select the appropriate surrogate decision maker when the patient lacks decision-making capacity;
- Understand the grounds on which surrogates should make decisions for patients who lack decision-making capacity." (page 29)

The issue of patients whose decision-making capacity may be in doubt is only a small part of a much wider problem of patient understanding. The GMC's duty "to give information to patients in a way they can understand" must be one of the most difficult to fulfil; and "inadequate explanation" is one of the most frequent forms of patient complaint. Moreover, the consequences for the quality of care tend to be underestimated. [SFM80]'s research review suggested that the way doctors interact with their patients affects the adequacy of clinical interviews, patient compliance and satisfaction and responses to potentially distressing medical and surgical procedures, all of which may in turn affect long term outcomes for patients. They also presented evidence of the consequences for patients of ineffective communication skills in junior doctors. A later study by [T...85] of doctor-patient communication in general practice found that in "as many as one in every two consultations patients could not recall all the key points..., could not make correct sense of them, or were not committed to them" (pp 167-178)

> "Because doctors did not know the details of what patients were thinking, the information they did give could not relate, in any precise or considered way, to the ideas patients themselves possessed. In short, there was little dialogue and little sharing of ideas. In consequence, doctors could have no way of knowing whether the information they offered was being understood 'correctly' or not. Equally, patients could have no way knowing whether their understanding of what doctors said was 'correct'." (p 205)

At a more general level, and particularly in community settings and clinics, it is important for the doctor to find out what is worrying the patient and why they have come for a consultation in order to frame the health problem in an appropriate way for progressing the situation. The doctor's problem may not be the same as that of the patient ([McW85]).

Another aspect is communication with relatives and/or friends, when there is patient consent to informing them. Patients benefit from the opportunity to talk things through with family and/or friends, who also need to be well informed. Moreover drug compliance, reporting of side-effects, and changes to diet and/or lifestyle are more likely to be sustained when there is family understanding and support. In many cases continuing communication with both patients and families will need to be part of the care plan. Nurses and therapists will often have good opportunities for this; so the whole care group need to be well briefed and to know who is responsible for talking to whom, and for checking up later on how much has been understood. In some cases, especially where there are psychological aspects to be investigated, the problem may not be confined to the person who first sees the doctor. Both the framing of the problem and the response to it may involve skilful communication with several people.

Health promotion and preventive medicine also require good communication skills with patients, relatives and the wider community, as well as the disposition to seek and take opportunities to talk about the relevant issues. The administration and management of communication is also an important area of competence, and this does not mean doing everything oneself. The GMC publishes guidance on serious communicable diseases and fitness to drive a vehicle; but there is little reference to the detection of risk-bearing psychological problems by doctors outside general practice and psychiatry. Given the increasing emphasis on preventive medicine policies, it would seem reasonable to suppose that some of the competencies required for membership of the Faculty of Public Health Medicine ([KD97]) might also be needed by other doctors, e.g. good understanding of local variations in health problems, epidemiological approaches to the assessment of health care needs, evaluation of health care policies.

2.4 Management Competence

Teamwork, management and leadership skills are needed in a wide range of contexts: within the relevant care groups and the medical practice or firm; and in order to perform duties needed for the effective and efficient functioning of the health care organisation in the interests of its patients, employees and other stakeholders. Interprofessional communication, team working skills, record-keeping are frequently cited examples. So are teaching and training; but not appraising, mentoring and giving feedback. Practice evaluation and quality improvement skills cannot be taken for granted but require systematic development ([EM96]). The FCIM Task Force listed 13 competencies for the Management of the Quality of Health Care, mostly methodological (e.g. measuring patient satisfaction, measures of severity of illness and comorbidity, knowing methods for evaluating the effectiveness and efficiency of one's practice patterns) but also including.

> "Know how to lead a health care team that is trying to improve the quality of its services (understand team behaviour, working with a team, and reshaping a team)." (page 59)

The research on changing primary health care points to the need for multiple interventions, which combine the development of competence through CME with interventions aimed at changing aspects of practice management (see Chapter 6). Often noted gaps between individual or team competence and performance may also be attributable to lack of management competence within the team or at higher levels. This may be highlighted by the development of clinical governance for which the link between quality and improvement and management responsibility will be critical.

Finally, we come to the competence to manage one's job as a whole. This includes determining priorities, implementing professional values, improving the quality of one's personal practice, self-evaluation for all one's roles and duties, monitoring one's own health and stress-level and taking appropriate action. The importance of competence at a whole job level is seen by many as necessary for ensuring that competence in different aspects of the job is translated into performance. An emerging issue in this area arises from a greater wish by doctors to protect their personal lives and avoid high stress and burn-out by limiting their hours of work ([YLSH93]): doctors too have rights.

However, this can be interpreted as unprofessional by "the older generation". Apart from the problem of overall staffing levels, the central issue is often continuity of care. This is turn depends on the ability of the whole health care team to organise itself in a way that ensures such continuity. The notion of team competence and organisational competence are being increasingly used in the private sector. While this goes beyond the brief of this report, it indicates the extent to which interprofessional work is likely to be expected of doctors in the future and hence the attributes of competence required for it.

2.5 Judgement

The term **judgement** is less controversial than competence for most doctors, because it conveys a sense of expertise. However, some gynaecologists now avoid the term 'clinical judgement' because it has acquired a connotation of paternalistic devaluing of women's choice. One meaning of the term is legalistic: a court or a disciplinary hearing makes a judgement; and this meaning is often extended to cover the process of assessment, where examiners also are described as making judgements. More frequently, 'judgement' is a term used to describe the highest level(s) of expertise. According to whether they use a broad or narrow definition of competence, they will define judgement either as an advanced level of competence or as that area of expertise which goes beyond competence. The most salient attributes of judgement reported in our consultative interviews concerned making holistic and balanced decisions in situations of uncertainty and complexity. More specifically, descriptions of bad judgement included:

- not integrating all the data
- attending to fine detail, "the small print", but missing the big picture
- choosing an inappropriate management plan
- overvigorous intervention
- not taking into account conflict between different therapies
- being competent in all the important aspects of the case, but failing to make a sensible decision on the case as a whole
- making decisions on current evidence alone and disregarding the record.

Good judgement, by implication, is the opposite of bad judgement, but is not always described in the same terms. For example, good judgement could mean:

- discerning the key features of a patient's problem in a more complex way
- going beyond the guidelines
- intuitive but rationally checked out expertise
- making small approximate decisions and readjusting

- being prepared to do nothing.

Situations where judgement (hopefully good) was called for included:

- decisions based on fuzzy logic in situations too complex to fully understand
- ill-defined situations which are complex diffuse and muddled
- high risk situations
- cases combining medical and psychological aspects
- deciding what to tell a patient and how to put it
- prescribing and adjusting drug "cocktails"
- deciding between maximally invasive and minimally invasive procedures (or doing neither)
- balancing cost and quality

Other aspects of judgement described in the literature concern:

- ethical issues
- the effect of a doctor's feelings and emotions about a particular case-these could concern the patient situation, their own behaviour, the actions of other health care professionals or the health care organisation ([FC98]).

Our consultations revealed the expectation that, given sufficient time, all postgraduate trainees would become at least minimally competent doctors. If they failed, this would normally be due to dispositional factors — a profound disregard for patients and/or an unwillingness to evaluate their own practice and acknowledge where improvement is needed. The self-evaluation problem might concern not just a doctor's personal practice but also his or her role in the health care team and the practitioner community. Difficulties in moving from completion of training to the role of consultant were often attributed to overconfidence, failure to recognise what further learning was needed and a failure to fit into the new context and grow into the new role. This also was sometimes described as bad professional judgement. Whether clinical or more broadly professional, judgement takes time to acquire and its development cannot be guaranteed by good training.

2.6 Summary

Competence is defined as "the ability to perform the roles and tasks required by one's job to the expected standard". This definition does not apply only to the intended outcomes of basic and specialist training. Throughout postgraduate education doctors have direct responsibilities for providing patient care and their current competence has to meet this service expectation as it changes with increasing experience and seniority. They are also seeking to expand that competence to meet anticipated future responsibilities; and this dual obligation presents a considerable challenge. Their success in meeting this challenge is significantly affected by the balance between the demands of their job and the opportunities presented by their learning programme.

Further insight into the nature of this challenge is provided by recognising the distinction between **competence**, what a doctor can do, and **performance**, what a doctor does do. There are many reasons why a doctor's performance may not reach the level of their competence. Some of these are associated with personal factors such as disposition, attitude and life-events outside the job. Some are associated with contextual factors such as workload, working conditions, levels of support and unexpected bursts of emergency work.

The difficult task of specifying competence is discussed with reference to finding the appropriate level of detail, capturing the essence of expertise, recognising the integrated nature of performance, and covering the full breadth of the doctor's job today. The evidence feeding into this process tends to come from committees of experienced physicians who also have a record of contribution to postgraduate education. Contributions from research are rare but might be useful if based on studying doctors' responsibilities and relationships as well as their clinical experience. These have been clearly addressed by recent publications of the General Medical Council; and an excellent example of a complete analysis of the competency goals of residency programmes and planning appropriate ways to achieve them is provided by the US Federated Council for Internal Medicine Task Force. The significance of communicative competence and management competence in particular is discussed in separate sections.

The chapter concludes with a brief summary of interviewee responses to questions concerning salient attributes of **judgement** and situations where judgement is required. Though not providing an agreed definition of **judgement**, the analysis highlights aspects of a doctor's work where feedback on performance, case discussion and possibly even coaching are especially important during postgraduate medical education.

Chapter 3

Theories of Expertise

Until recently, research into professional education suffered from two major weaknesses: the knowledge-base of professionals was conceived largely in terms of formal, published knowledge; and research into learning was also focused primarily on formal educational contexts.

Practitioners were usually aware of these weaknesses and expressed their concerns by using the currency of experience (months in a job, or number of cases seen) to complement the system of specialist examinations. The outcomes of learning were expressed in terms of competence (a generic term signifying capability to take independent action with only occasional referral to others) or judgement (a term conveying a mysterious ability to make wise and effective decisions in situations of considerable complexity and uncertainty). There was little attempt to ascertain how competence and judgement are acquired or how the learning of this practical knowledge might be facilitated.

So while educational practice continued to give central place to the doctor's practical experience, the relationship between this experience and the growth of competence and judgement remained mysterious.

Since the early 1970s, researchers have begun to address the hitherto uncharted territory from a number of perspectives. For a review of the evolution of the notion of clinical competence, see [Maa90]. These approaches to understanding the nature of medical competence and judgement can be characterised into three main types and one minor type:

- Approaches that focus on what experts know and how they know it, their memory structures, their reasoning and how all these differ from non-experts. Typically this approach makes use of (psychology) laboratory-based studies.

  A stimulus was the advent of the computer as an aid to representing knowledge and an interactive guide to decision-making. The attempts to design expert systems for this purpose helped to create new theories about how experts (and also novices) think and greater awareness of other modes of cognition than computation and logical deduction. These include concepts much as 'schemas' (accustomed ways of thinking) 'scripts' or 'protocols' (blueprints for a series of sequential actions) and 'mental models' (how people perceive a situation habitually), all of which have proved fruitful in describing medical expertise and its acquisition. Other

aspects of cognitive psychology now perceived as relevant include studies of bias in human judgement and of the use of different kinds of memory (semantic, episodic, procedural). These can all be seen as steps along the path of modelling experts' ways of appreciating the medical situation 'as a whole' — a criticism of some early expert systems approaches ([War86]).

- Approaches that focus on how experts behave in real settings, what they pay attention to, how they allocate time and other resources, their reasoning and how all these differ from non-experts operating in similar situations. Typically this approach is not conducted in the (psychology) laboratory.

  While cognitive psychologists have concentrated mainly on a combination of laboratory based studies and dialogues with experts, other social scientists have adopted a more ethnographic approach to studying decision-making in naturalistic settings. This supports the cognitive psychologists' view that reasoning is mainly schema-driven with decision-makers creating causal models of the situation or, more often, matching it to models they have used before. But it also draws attention to characteristics of naturalistic settings such as time constraints, high stakes, ill-structured problems, incomplete information, shifting goals and many participants. This makes the ability to assess situations rapidly a distinguishing feature of expertise, and a readiness to settle for 'good enough' rather than optimal solutions. Time constraints also lead to early selection of a highly likely option and checking it out rather than generating several options at a time. Moreover decisions are often multiple rather than single, as people "think a little, act a little, evaluate the outcomes then think and act some more".

- Approaches that focus on what an ideal expert might be expected to do by reference to the best empirical evidence available (e.g. evidence-based medicine), or to idealised models of decision-making under uncertainty or to computerised decision support systems.

  For example, the growing number of epidemiological studies led to techniques for making better use of the probabilistic evidence they provided; and this approach, now called decision analysis was extended to incorporate evidence of clients' evaluations of the desirability of various possible outcomes.

- Approaches that focus on experts' sense of their own expertise, their comparative self-confidence in their competence to deal effectively with the difficult decisions they have to make.

Each general approach has its strengths and weaknesses and each illuminates different facets of medical expertise.

Issues that tend to get downplayed in all the above approaches include interpersonal and communicative skills, manual dexterity and hand-eye coordination, and overall issues of patient management and care.

3.1 What Experts Know

3.1.1 Decision Making Skills

Much of the psychological literature on medical expertise has concentrated on the important issue of decision-making. Many of the studies are conducted in the laboratory and involve presenting descriptions of cases to doctors of varying degrees of expertise in order to observe what aspects of the case are paid attention to, what inferences are drawn at various stages, what hypotheses generated and rejected, what overall conclusions derived and what aspects of the case are later remembered. These kinds of study are used to delineate the evolution from novice to expert in terms of how various kinds of medical knowledge are organized in memory, how this knowledge is accessed and how it influences decision-making. Most studies are based in areas of internal medicine. Studies of radiological expertise additionally examine how radiologists view images, and how their perceptual scanning changes with increasing expertise. Relatively few studies examine the knowledge underpinning surgical expertise.

This literature makes it clear that there is an evolution of both knowledge structure and diagnostic skill from novices through intermediates to experts. This is an evolution both in terms of how much is known but, more importantly, in the organization and structure of what is known and how medical problems are represented ([CBC98]). The categories "novice", "advanced-novice", "semi-expert", "sub-expert", "expert" and even "super-expert" ([REM98]) are not well defined and the reported discontinuities of knowledge organization between these categories can make it look as if there are "stage-like phenomena" at work. Indeed, there is disagreement among researchers as to whether there are distinct stages or whether these are (to some extent) an artifact of the methodology (see e.g., [PG91]). Whatever the truth of this issue, most researchers agree that the later 'stages' of the development of expertise are characterised not so much by further increases in knowledge of pathophysiology but by changes in the organisation of that knowledge to make it more readily and rapidly available. There is also agreement on the importance of decision-making with real cases within working medical settings as both driving that change and shaping its nature. One of the side-effects of this is that an expert's knowledge is highly personal and depends very strongly on the particular cases which that expert has encountered ([GM88]).

Novice medical decision-making can be characterised as working largely from scientific first principles as opposed to clinical principles. They also reason largely 'backward' from hypotheses (though see [APP93]), as opposed to 'forward' from the data. They are able to construct only a limited set of hypotheses, are not able to evaluate competing hypotheses well and are not able to deal properly with apparently inconsistent data. Although working within a different paradigm (phenomenography rather than cognitive science), [RWC89] capture the essence of the novice approach:

> "The distinction reveals itself in this material in the form of a contrast between a focus on specific symptoms and signs, or short links between causes and effects, and the use of structuring principles to systematize the data and relate them to previous knowledge. In relation to the very complex material constituting a disgnostic task, an atomistic (ordering) approach does not mean that analysis, interpretation and organization of the material

> are completely absent. These learners clearly display the rational use of these skills; many apparently have a knowledge base that is large enough to enable them to deploy them in a way that a more experienced learner would. Lack of procedural skill and preclinical knowledge is not an explanation for their use of atomistic approaches. What appears to be missing in the attempts of the students ...is competence in representing the problem appropriately, so that its inherent structure is maintained." ([RWC89], pages 113–114)

By contrast expert medical decision makers work with highly structured knowledge ([Bor91]) that provides various kinds of shortcut to the small set of hypotheses that need to be considered in any situation. While the fact that experts' knowledge is highly structured is relatively uncontentious, the exact nature of these structures is a matter of debate (e.g. whether more based on generalisations or whether more based on accumulated instances). For a detailed, critical review of Prototype Frameworks, Instance-based Frameworks and Semantic-networks, Schema and Script models, see [CRN96]. For an example of the educational repercussions of the distinction between instance-based theories and abstraction-based theories, see [PSA96]. The distinction here is between the expert concept of a disease being internally represented largely in terms of particularly telling example cases, as compared to the disease concept being represented as an abstract generalisation that glosses details of individual cases. Whichever theory is correct, [PSA96] point out that both theories support the finding that more typical cases are more likely to be correctly classified. They suggest that teaching and assessment could be refined to take specific account of the typicality of the cases that they teach/test.

Experts are data driven and don't appear to work directly from scientific first principles so much as from an "illness scriptillness script" ([SNB90]) that encapsulates various levels of knowledge (including, at base, the scientific) in a schema (conceptual structure) associated with a particular pathology. When presented with a new case experts rapidly home in on a number of "critical cuescritical cues" (see e.g., [CP87]) and key features (for assessment methods based on this, see [PBA95]) that guide them to consider a small set of possible hypotheses ([KPM98]). Discriminating between the hypotheses in this small set is partly dependent on the "relative distinctiveness of [the] competing classes" ([PSM90]).

In many ways this rapid homing in process is largely unconscious, though the end result is reflected on and regulated by conscious processes ([Bor94]). Experts are also strongly guided by "enabling conditionsenabling conditions", i.e. crucial factors in the patient data or clinical history. Experts also have an excellent memory for the relevant aspects ([PGF86]) of exceptional individual cases that they have seen and use these in dealing with new cases, though such memories can also be a source of distorted learning (Featherstone, 1984).

Intermediates largely fall between novices and experts, but sometimes perform worse than novices on some laboratory-based tasks ([LRF...88]), and sometimes perform better than experts on other laboratory-based tasks (i.e. because they have paid attention to different aspects of the case description from experts). However this "intermediate" effect is not always found ([vdWSB98])

Development of Expertise

The literature in this area is large, but a useful accessible starting point is the paper by [SB93]. The authors set out a model for the development of expertise that marries together the influence of formal education in basic scientific knowledge as well as practical experience of dealing with cases. They summarise the evidence for a *staged* account of the development of expertise as opposed to an incremental approach i.e. that what experts know is organized differently from what novices know and is not just a matter of knowing more than novices. In particular, they argue that the a novice's knowledge is re-organized a number of times on the way to expertise. The paper describes the different kinds of knowledge structure that evolve with time and outlines the evidence for the existence of "illness scripts". Illness Scripts are stereotypical accounts of the enabling conditions, predisposing factors, special conditions, causation and consequences (e.g. complaints, signs and symptoms) for a particular disease.
The authors argue for the relatively greater importance of instances of past cases encountered compared to basic scientific knowledge in the expert diagnostic reasoning process, i.e. that it's the experience of dealing with actual cases rather than scientific "bookwork" that makes the most difference. While the details of the exact forms of the knowledge of novices, intermediates and experts are beyond the scope of this report, the main issue is that knowledge built up in the later stages through exposure to actual cases has the strongest effect of how new cases are dealt with. Only where this is insufficient do experts fall back on the more basic scientific principles that they were earlier exposed to.

> First, "it assumes that the development of expertise can be described as the progression through a series of transitory phases. Second, knowledge acquired during different phases of expertise development has a distinctly different organization, earlier forms tending to be organized in causal networks, more recent forms being structured as scripts. Third, it is assumed that knowledge acquired in different phases form layers in memory through a sedimentation process. Fourth, these knowledge sediments, although usually not applied any more in subsequent phases in the development of expertise, remain available for use when more recently acquired structures fail in producing an adequate representation of a clinical problem. Fifth, episodic traces of clinical problems previously analyzed seem to be extensively used in the representation and solution of new cases."
>
> ([SB93], pages 217–218)

The above studies can, in principle, be criticised on two grounds. First, their conception of the nature of medical expertise is too narrow and focuses on decision-making to the exclusion of other aspects of patient management and care. Second, even within the area of decision-making the focus is on idealized, laboratory-based decision-making from descriptions of cases rather than real decision-making associated with actual cases as conducted in working medical settings. For example, the decision-making is affected by the context in which it occurs, and a skill learned in one context does not necessarily

transfer readily to a different context (see e.g., [Gru97]). However it turns out (see below) that the analysis of expertise in terms of critical cues, small worlds and illness scripts is not that far removed from analyses based in the naturalistic decision-making paradigm. Indeed, [PKA95] provide a detailed account that tries to reconcile studies of medical knowledge and expertise (such as cited in this section) with a more "situated" view that recognizes the important role of context, of artifacts (such as decision-support systems) and of the collaborative nature of much medical decision making.

The following three short sections focus on expertise with respect to three special areas of skill — the perceptual (e.g. radiology), the psychomotor (e.g. surgery), and the communicative (all aspects of medicine).

3.1.2 Perceptual Skills

Most areas of medicine depend crucially on the doctor seeing what needs to be seen. However, while decision-making has been relatively well researched, medical perception has received rather less attention. An exception is radiology.

[NCBB92] review 46 studies in the area of visual diagnosis. They examine *inter alia* the effects of prior and concurrent information on radiologic diagnosis, the interplay between perception and analysis, and the question "are good diagnosticians born or made?". For example, on the issue of prior and concurrent information, they describe studies that showed that providing a tentative diagnosis increased true-positive rates of detection with a small increase in false-positive (but see below for studies that post-date this review and which indicate the dangers of prior information). They offer the following conclusion about the educational implications:

> "If one simply accepts that visual diagnosis does have two identifiable, although not entirely separable, components, it is evident that educational strategies directed at perception and cognition are very different. Perceptual skill is unlikely to be enhanced by any elaboration of rules or high-level processing of features lists or causal mechanisms, although this may well enhance cognitive processing. Rather, perception, with its rapid and *gestaltist* aspects, is only likely to improve from exposure to many carefully chosen prototypical examples and variations on the same theme." ([NCBB92], page S82)

In general the literature on radiological decision-making largely concurs with the decision-making literature described earlier but it also takes special account of perceptual processes and the way they interact with problem-solving ([Rog96]). One issue is the way that some aspects of perceptual skill may be acquired largely unconsciously and in a way that does not require their verbal articulation ([LHB88]).

With respect to radiology novices are slower and less efficient in scanning images ([NKLT96]), are less able to identify the 3D position of abnormalities and less able to identify the physical extent of the abnormality.

Expert radiologists are able to identify much of the abnormality in an image very quickly (an initial gestalt view) and this is followed by a more deliberative perceptual analysis.

They are better at identifying the 3D position of the image (i.e. responding to "localisation cues") and also better at identifying the physical extent of the abnormality ([LRF...88]). Experts have a better appreciation for the range of normality and have a propensity to pay attention and to recall abnormal cases better than normal ones ([MWJS88]). [REM98] distinguish "super-expert" from "expert" radiologists. In particular they found that super-experts were more alert to less salient factors in the images and more likely to construct an understanding that took all the factors into account. [REM98] suggest that this is partly due to their super-experts being exposed to more difficult cases and also because, as researchers, they would be used to devoting "conscious effort making their results explicit and publishing in scientific journals" (page 539).

All groups are sensitive to the effect of other information about the patient on what they see and the way in which they integrate visual and written information ([NBC...96, NBR...96]).

Consulting information about the patient prior to viewing the images affects not only what they see but also what they diagnose and therefore recommend ([BNC93]). A similar effect is also reported for ECG interpretation ([HNB96]).

Development of Radiological Expertise

[LRF...88] is a useful paper in this area. The authors describe a series of experiments in which they showed standard posterior-anterior thoracic radiographs to novices, intermediates and experts and invited them to describe what they saw. Their account provides both quantitative and qualitative data, and examines the interplay between cognitive and perceptual processing.
They note:

- the speed and accuracy with which experts build a mental representation of the abnormal anatomy shown in the radiographs.
- the speed with which experts invoke a likely schema to explain what they are seeing, and the way that such a schema subsequently guides both their perception and their reasoning.
- the flexibility that experts exhibit to fine tune a schema to make it fit the findings as well as their ability to make finer discriminations than novices.
- the fact that experts "see things differently" from novices; for example, experts regarded, and traced as abnormal, a larger area from a collapsed lung film than either novices or intermediates.
- experts reasoned "opportunistically" incorporating new pieces of data into their diagnostic decisions.

[LRF...88]

3.1.3 Psychomotor Skills

As with perceptual skills, effective hand-eye coordination and the ability to carry out procedures is of great importance all across the field. For an extensive theoretical review of the issues of learning and retention of motor skills in general, see [Ada87]. Here we will concentrate on surgical skills. For a much more domain-specific account of the problems of learning and teaching surgical skills (including the use of 'Craft Workshops'), see [Kir96].

The evolution of surgical skills, where they are different from the decision-making skills already discussed, has not received anything like the same degree of attention. So, for example, we tend not to find studies contrasting experts and novices.

[Bar87] provides a useful overview of some of the issues underpinning the development and the teaching of surgical skills. He mentions (by reference to [Kop71]) determinants of surgical skills including "speed, accuracy, economy of effort, and adaptability". Pointing to [LF84], he reminds us that psychomotor development has to take place alongside the cognitive and affective. He also characterises the surgeon as a problem-solver who in the preoperative phase must develop "a systematic technique for selecting the most appropriate procedure"(page 424), and in the postoperative phase must develop the ability to reflect on practice effectively. The literature on "naturalistic decision-making" (see below) characterises expertise in terms of problem recognition, and gives support to the idea of the expert (e.g. a surgeon) as someone who anticipates and plans for potential problems, rather than simply reacting to them when they occur, see e.g., [XMD97] in the field of anesthesiology.

By reference to the literature on motor skills development in sports education, [Bar87], who characterises surgery as the "the ultimate body contact sport", emphasises two principles.

The first principle is to prevent learners developing faulty initial habits which are then very hard to unlearn. Given that much is learned in apprenticeship mode, poor role models can have a crucial effect.

The second principle is that "skill retention correlates with the level of initial proficiency and not with practice". While this cannot be used as an argument against practice, it does emphasise that getting the skill right early in training is very important. Indeed surgeons do get better with practice (see e.g., [BBG97]), but this is hardly surprising.

Barnes goes on to argue for and to describe a number of microsurgical training laboratories through which guided supervision and practice can be provided.

An example of a non-surgical focus on psychomotor skills is provided by [Kov97] who describes a methodology (in the area of Advanced Trauma Life Support — ATLS) for developing procedure skills that has the following stages:

1. Conceptualization
2. Visualization
3. Verbalization
4. Practice-subcomponent, linkage, continuous

5. Correction and Reinforcement
6. Skill Mastery
7. Autonomy

He emphasises the key point, in relation to item 5, that:

> "... knowledge of results is required to learn, *correct, and improve the performance of motor action.* This principle is frequently violated in medicine where so many procedures performed by house staff *go unobserved.*" (our emphasis) ([Kov97], page 389)

We return to this issue of supervision and feedback in Chapter 4 on Learning in Clinical Settings.

A similar focus on initial accuracy followed by later practice for speed is described by [SJHC97] in their study of learning curves for fibreoptic nasotracheal intubation. They found that following initial training, it took on average 18 practices (under instruction) with real patients to reach a 70% criterion of completing the intubation in less than a minute. Not all procedures have to be practiced with real patients, as is shown by [DBSF98] using a laparoscopic simulator, and more generally by [HBC95] in their review of ten years experience of a psychomotor skills laboratory.

3.1.4 Communicative Skills

The significance of communication in different zones of medical activity was discussed in Chapter 2. Two groups have been given most attention in the literature: medical students, though not with the priority many authors would like; and General Practitioners, for whom it has been a major concern. Theoretical discussion about communication has largely focused on Medical Interviews in GP surgeries or clinics, except for the specialist area of psychiatry which we are not attempting to tackle. From a general medical perspective, [HHM...75] showed that in 82% of general practice consultations the diagnosis reached could have been made on the basis of history taking alone, without either a physical examination or any laboratory test. Hence communication skill was required for eliciting most of the important evidence. [Bal57] suggested that most problems have a psychological element which needs to be explored; and that this was the most significant aspect of at least 25% of cases. Given that even physical symptoms of importance for diagnosis may not be volunteered in the early stages of a consultation, Balint's notion of a "deeper diagnosis" has acquired some credibility, and this has led to the use of a continuum from doctor-centred to patient-centred as perhaps the main dimension for analysing GP consultations. [BL76]'s study of over 2000 GP consultations used a 7 point continuum and identified about 20 doctor-centred, 20 patient-centred and 8 negative behaviours as a framework for analysing interviews. They also distinguished six phases to the interview (see box below) and showed that consultations were particularly likely to go wrong if there were shortcomings in Phase 2 or Phase 4. Their conclusion was that a more patient-centred approach was more likely to elicit important information about the psychological and physical symptoms:

Doctor-Patient Consultation

1. The doctor establishes a relationship with the patient.
2. The doctor attempts to discover, or actually discovers, the reasons for the patient's attendance.
3. The doctor conducts a verbal or physical examination, or both.
4. The doctor, or the doctor and the patient, or the patient (in that order of probability) consider the condition.
5. The doctor, and occasionally the patient, details treatment, or further investigation.
6. The consultation is terminated, usually by the doctor.

[BL76]'s six phases of a doctor-patient consultation (as described by [P...84]).

Sociologists have drawn attention to the role expectations of doctors and patients which facilitate or constrain communication, including the effects of social class ([Bai76, Bai77, PB80]); [Kle80] noted ritualistic parallels between traditional and 'Western' healers. Several anthropologists have distinguished between diseases as labels given by doctors, and illness as a broader concept defined in terms of the patient ([Hel81]). This includes the response of the patient to a problem, how it affects the patient's behaviour or relationship, the patient's past experiences of illness and the meaning she gives to that experience ([P...84]). [L...86] rightly emphasise integrating a patient-centred approach with the doctor's concern for differential diagnosis, as a purely Rogerian approach does not necessarily elicit more relevant information ([BS85]).

Hence the view that the doctor's understanding of the problem must include an appreciation of the patient's own understanding, whether or not she perceives it to be accurate. Without such understanding patients are unlikely to understand what the doctor tells them or comply with this advice ([Bec79]). This Health Belief Model, described in detail by [P...84], is also much used in Preventive Medicine.

Empirical evidence in support of these theories is reviewed by [P...84] who concluded that:

> "Patients are more satisfied when the doctor discovers and deals with patients concerns and expectations; communicates warmth, interest, and concern about the patient, volunteers a lot of information; and explains matters to the patient in terms that are understood."

A comprehensive study by [BRP91] of 550 consultations by 127 physicians in the US (90% in internal medicine, 35% residents) concluded that patients were most satisfied by interviews that "encourage them to talk about psychosocial issues in an atmosphere that is characterised by the absence of physicians dominance". [Bec79]'s Becker's (1979)

review of compliance demonstrated that patients comply better when they believe they can have control over their health and when the advice given is consistent with their own health beliefs; and [KN72] showed that mothers leaving a paediatric clinic "highly satisfied" (40%) were three times more likely to follow the doctor's advice fully than those who were "highly dissatisfied" (13%). The main reasons for dissatisfaction were unfriendly behaviour, and the lack of information about the nature or cause of their child's illness. At a more general level, [FF88]'s review also concludes that most patients want more information than they are given.

The Toronto consensus statement issued by a meeting of researchers in this field ([SBS...91]) reviewed evidence about the significance of doctor-patient communication, then advised on the teaching of communication skills as follows:

> "To become effective communicators, physicians must master a defined body of knowledge, skills, and attitudes. Clinical communication skills do not reliably improve from mere experience. Examples of relevant areas of knowledge are psychiatry in relation to medicine (for example, diagnostic clues to depression, anxiety, somatisation problems) and the structure and functions of medical interview are those of data gathering, forming and maintaining relationships, dealing with difficult issues (such as sexual history, breaking bad news, HIV), and imparting information; therapeutic skills and strategies are also necessary. These skills can be defined with behavioural criteria and can be reliably taught and assessed. Helpful attitudes include a belief in the importance of a biopsychosocial perspective. A physician's personal growth and self awareness are essential bases of effective communication." ([SBS...91], pages 1385–6)

Cognitive science concepts similar to those used in theories of diagnostic expertise have also been applied to doctor-patient consultations. [TW86, TW87] analyse transcripts in detail, but instead of using the wide repertoire of behaviours identified by [BL76], they use the concepts of 'register', 'frame', 'schema' and 'script'. Their data set is a series of videotaped conversations in five different settings, involving various family members and medical professionals in a single paediatric case of cerebral palsy. The videotapes were also intended for training paediatric residents. In one social encounter involving a paediatrician, mother and child, the following phenomena were noted:

- Three Linguistic Registers were used by the doctor — "motherese" to the child, medical language in a flat intonation for the residents watching the video, and ordinary conversation with the mother.
- Three Competing Frames affected their behaviour — a physical examination of the child, entertaining her with a playful, teasing approach; ignoring the video while still explaining what she was doing for a video audience; and conducting a consultation with the mother.
- Competing Schemata for understanding the symptoms being discussed. The mother was concerned about her child's noisy breathing and a skin condition, neither of which were significant for the paediatrician monitoring the progress of the cerebral palsy.

- An illness script for cerebral palsy which dominated the doctor's thinking. The breathing and skin condition did not figure in the script, but the colour of the child's skin at birth (several years previously) was in the script. This was agreed as having been "yellowish-white" with the parent coordinator/history taker, though subsequently it became clear that this negotiated "fact" did not match the parents memory!

Having such a script is probably what enabled the paediatrician to navigate her way through this complex situation, though it also brought disadvantages. Later they decided that it would be better for parents to watch through a one-way mirror while the child is being examined, thus separating over time the examination and the consultation. [KH95] have explored the practical implications of the overt incorporation of cognitive schemata and script development into the teaching of communication skills.

There is also evidence that doctors and dentists may use their expertise and other conversational methods to persuade patients to accept treatments favoured by themselves thus, not necessarily with intention, diminishing the patient's power to choose. This complex self-awareness problem is most easily handled by analysis of recorded conversation in a safe context. [Bal57] refers to it as the "apostolic function" of the doctor, and [And86] has produced similar evidence for dentists.

Another aspect of communication noted but not deeply studied among doctors or health care groups is teamwork. However, there in an extensive non-medical literature on team performance and training ([SS92]) which needs addressing. One conclusion, for example, is that individual skill training will enhance overall team performance in situations of low task complexity; but in situations of high task complexity and high task organisation effective communication and coordination among team members are vitally important ([S...92]).

3.2 How Experts Do Decision Making

A different branch of the expertise literature, Naturalistic Decision Making, has focussed its attention on what experts actually do when they work. Much is made of the difference in emphasis compared to the laboratory studies described earlier. Many of the studies focus on military and industrial settings (e.g. the control of complex processes such as in power stations), but a number have also examined practical medical decision-making (e.g. in anaesthesiology).

> "It is therefore necessary to move away from the traditional view of diagnosis as a cognitive problem-solving activity which establishes a complete scientific explanation of the cause of the malfunction before the operator chooses a course of action. Instead, a theory of diagnosis must be a theory of optimal behaviour, in which some control actions may be performed prior to information gathering, and information gathering may be curtailed before a complete picture has been established in the interest of maximizing utility." ([HAB95]) (page 22)

The issues which distinguish this approach to understanding expertise are summarised by ([ZK97], page 5) as:

- Ill-structured problems, not artificial well-structured problems.
- Uncertain, dynamic environments, not static, simulated situations.
- Shifting, ill-defined, or competing goals, not clear and stable goals.
- Action/feedback loops, not one-shot decisions.
- Time stress, as opposed to ample time for tasks.
- High stakes, not decisions devoid of true consequences for the decision maker.
- Multiple players, as opposed to individual decision-making.
- Organizational goals and norms, as opposed to decision-making in a vacuum.

Much of literature is concerned with the issue that experts have become attuned to the kinds of problem that they are likely to face and have developed strategies to deal with them. The issue then becomes one of recognizing the nature of the problem ("Recognition-Primed Decision Makeing", [Kle89]) It also involves monitoring the effectiveness of a largely pre-formulated way of dealing with it. Pre-formulation arises from the need to save time and make rapid decisions (not necessarily under stress). In a way this is reminiscent of the "illness script" mentioned above, but it also brings into play the dynamic factors in the above list (e.g. time stress).

[Lip93] provides an overview of various emerging theories of the decision-making of expert individuals in naturalistic settings. He identifies a number of common themes.

- Diversity of form: it is clear that different kinds of real world decision-making work in different ways (e.g. contrast the way a jury has to reconstruct the most plausible story that accounts for the events they have been told about, with the generals' decision to attack or not potentially hostile aircraft as seen on a radar screen). This suggests that there may be a similar diversity of decision-making expertise in medicine. For example, a gynaecologist cooperatively deciding with a woman about whether or not to have an abortion is likely to use a different kind of process from an anaesthetist deciding how to deal with an emergency during an operation.
- Situation assessment: all the models involve a process of identifying the nature of the problem, deciding what's relevant and what is irrelevant and classifying the problem into some pre-existing category. In many ways this aspect of naturalistic decision-making offers a similar account to that of the cognitive scientists who speak in terms of critical cues, illness scripts and small worlds. They offer a similar emphasis on matching the current case to critical former cases as opposed to reasoning from basic (scientific) principles.
- Use of mental imagery: rather than seeing the decision-maker as (mathematically) weighing the costs and benefits of different courses of action, they see him as building a mental model of the situation. Again the notion of an illness script is not dissimilar.

- Context dependence: because decision-making is seen as operating within a context rather than as a disembodied, ethereal activity, factors in that context matter. Familiarity of the context, stress, time dependence among others all influence what factors are taken into account and how they are taken into account.
- Dynamic processes: again this backs up the view from cognitive science that decision-making is a complex process that to some extent can be decomposed e.g. into recognition and then action selection, or into cycles of intuitive mixed with analytic decision-making.
- Description based prescription: while we cannot argue that expert decision makers never make mistakes or are never open to biases, it is important to realise that there may be excellent reasons for the disparity between the way that decisions are actually made and the way some might prefer them to be made.

Dynamic Decision Making

A useful entry point to this literature from a medical point of view is the paper by [Gab92] who describes the issue of dynamic decision-makingdynamic decision-making in the field of anesthesiology. He stresses the dynamic issues: (i) "The pace of decision is determined externally ... events may occur frequently ... Some events cannot be avoided"; (ii) "The system is complicated and has many interconnected parts"; (iii) "There is uncertainty". Signals from instruments have to be interpreted and may be either weak or unreliable; and (iv) "There is risk". These factors underline the stressful nature of this work and the need for ways of assessing how decision-making in this area is affected by stress ([BJ97, BSJ98]).
Gaba, like Lipshitz above, reviews the work of the main investigators in the field. For example, in reviewing the work of [Woo90], he mentions dynamic decision-making biases such as (i) "cognitive tunnel vision" (where new data are coerced to fit a pre-existing and incorrect view of the situation); (ii) attending to surface issues rather than engaging with the underlying problem; and (iii) 'micawberism' — believing that everything will work out OK in the end, despite all the contra-indications.
Gaba goes on to describe some of his own experiments in the area and outlines a model of dynamic decision making. In simulator experiments (in realistic settings with realistic instrumentation, but using an intubation/thorax mannequin), he found large variations in performance among subjects and large variations across incidents. In agreement with the literature, he found that experts were better at anticipating problems and were more willing to "interact forcefully" with the surgeon.
Finally, Gaba argues that anaesthetists ought to be trained explicitly in crisis management in a way similar to the training given to pilots. That is, the training should explicitly address the way that stress, risk, complexity and lack of time can lead to decision-making biases, such as cognitive tunnel vision, *and* should train anesthetists in strategies to combat these biases.

([Gab92])

At a much more applied level, [Bog97] offers a brief introductory account, from an American perspective, of this approach to understanding the pressures on decision makers in

the area of healthcare. She reviews work that has examined time pressure, fatigue and stress as they impact on emergency surgery and anesthesiology: for example, an increase under stress in the reliance on technology by anesthesiologists. Other factors considered are the pressure to reduce costs, the variability in the quality of feedback and the interface with the many (complex) technologies in use, and the problems of shared responsibility (e.g. with surgeons) and communication amongst the team associated with the case.

3.2.1 The Sense of Expertise

Given the pressures under which medical expertise is typically exercised, it is important that experts develop self-confidence and a "sense of equanimity" ([YHA...98]). For example, [RBF...91] found that *non*-cognitive factors (such as conscientiousness and confidence) were strongly correlated with overall performance in anesthesiology departments. They urge that educators "reconsider the lack of emphasis historically placed on the noncognitive aspects of performance" (page 361). [YHA...98] found that giving students a chance to practice skills (such as history taking) via an assessment using stations in a situation where the students knew that they would not be 'failed' improved their self-confidence in the skills assessed. They also found that the sense of self-confidence was well-founded in that it was correlated with scores in the assessment. The educational lesson from this is that this kind of assessment instrument can be used to provide feedback to students about their progress, and knowing that you are making progress improves self-confidence.

Tracking changes in self-confidence over longer time periods is more problematic. [SKD96] compared responses using a critical incident technique between 18 doctors in their first 6 months of general practice with those same doctors 12 to 18 months later as they completed their advanced or mentor terms in the Royal Australian College of General Practitioners Training Program. In the interviews "doctors were asked to describe incidents and to identify skills either present or lacking, feelings and lessons learned". The researchers found that

> "increasing clinical practice ... resulted in an increase in the positive feelings associated with making a difficult diagnosis and dealing appropriately with more difficult management problems without immediate referral to a specialist.
>
> ...
>
> Doctors in the first interviews also reported feeling pressured for time, uncertain or anxious about possible missed diagnoses and inadequate management, and unsupported by supervisors and practice staff. Hardly any of these issues were mentioned in the follow up interview. Although doctors' levels of anxiety appeared less, they more frequently reported feelings of guilt over missed diagnoses or less than perfect management. With increased maturity the doctors appeared to become more 'tuned in' to the subtle interpersonal issues of the consultation and were more self critical regarding those episodes of miscommunication which result in a less than ideal consultation." ([SKD96], page S64)

3.3 What Experts Ought to Do

There are two overlapping approaches to the search for better and more consistent clinical decision-making. The first tends to focus on quantitative models of decision-making: for example, weighing evidence, estimating probabilities of outcomes and computing utility values. This approach naturally leads to the development of systematic (typically computer-based) methodologies to which the doctor can apply information and from which he or she can obtain advice that can be weighed against other evidence and test results. See the collection edited by [TS98] for a general overview of this approach.

A second approach focuses rather more on the provision of the best available evidence to assist doctors in their decision-making. Information Technology is involved here much more in its role as a repository and facilitator of access to information, than as a device to weigh information.

3.3.1 Clinical Decision Analysis

Decision Analysis

An excellent introduction to decision analysis is provided by [LT92]. They distinguish structured patient history taking from diagnostic systems and discuss the value of, and lack of general acceptability, of expert systems. They provide several examples of how to carry out a clinical decision analysis in terms of the treatment with the 'greatest expected utility'. They also are frank about the difficulties of carrying out a fully rigorous decision analysis in dealing with an individual patient. However they argue for its value in determining general treatment policy, in focusing research and in dealing with issues of communities as well as individuals.

[LT92]

[Dow93], in a book edited by [LH93], distinguishes between a number of questions that can be asked about clinical judgement and decision-making:

1. How are clinical judgements and decisions made? (To some extent this question is addressed in earlier parts of this chapter.)
2. How well are clinical judgements and decisions made?
3. How could they be made?
4. How well could they be made?
5. How should they be made?

Dowie concentrates on question 4, and reminds us that many studies have shown that in the area of clinical decision-making 'we could do better'. He goes on to argue that one route towards improvement is the development and use of system-aided judgement and decision-making.

In support of this view de Dombal (1993) reviews a number of studies that have demonstrated improved clinical decision-making performance, including a study that "showed

that when findings of detailed studies were made available to *inexperienced staff* performance improved in a number of hospitals" (emphasis ours). For example, initial diagnostic accuracy improved from 45.6% to 65.3% and post-investigative diagnostic accuracy improved from 57.9% to 74.2%. It has to be mentioned that this result is much more positive than that of [EFW...96], described in Chapter 7 on Information Technology.

The book edited by Llewelyn and Hopkins (1993) gives examples of how decision trees can be constructed, including methods to assess (numerically) the probability of every choice branch and methods to assess (numerically) the utility of each outcome state, as well as how to carry out a sensitivity analysis of the effect of varying the judgements. For instance, [HW86] provide a detailed example of the development of a decision rule for diagnosing pulmonary embolism using clinical findings.

A more detailed account of models of clinical decision-making is provided by [DE88]. This focuses on rational decision-making under uncertainty and its theoretical underpinnings. They include research on statistical approaches, including some which claim that in some decision-making contexts simple statistical models perform better than experts. They acknowledge that the relevance of these results to clinical decision-making tasks in realistic contexts is disputed:

> "Clinicians have great difficulty accepting that their expensive and complex technical knowledge does not necessarily guarantee that they can do better, in most clinical cases that come to them, than something as simple as "add up how many cues are in favour of each possible judgement and go with the highest score"." ([DE88], page 10)

Of course, part of the issue resides in the crucial word "most", in the sense that dealing as effectively with the unusual as with the usual, and knowing the difference between the two, marks out the difference between a simple rule of thumb and a more considered decision procedure.

The work of [MKK88] provides a direct empirical comparison between decision making of experts (using think-aloud protocols based on case descriptions) and decisions taken according decision analysis theory. They observed that

> "The experts did not formulate a global outline of their decision, but chained together a sequence of decisions based on available and incomplete information. Despite effective and efficient problem solving, the clinicians used numeric terms only as symbolic representations of likelihood, used limited information in choosing amongst alternatives, and dismissed the possibility that a less conventional strategy, empiric therapy, might yield equivalent outcome." ([MKK88], page 435)

It is argued that doctors often fail to carry out their decision-making in the hypothetico-deductive manner prescribed by these theories and use less rigorous strategies ([Mag92]) such as exploiting circumstantial evidence (see e.g., [BMF96]), or heuristics (see e.g., [Bor89]). In this case the question remains as to how decision support aids, that do work on a rigorous basis, can be exploited effectively by those who have in the end to

make, and bear the responsibility for, clinical decisions. Issues here include the degree to which the existence of the decision-making aid with its own particular methods and foci for information gathering about the case adds its own biases to the decision-making process (see e.g., [KP98]).

3.3.2 Evidence-Based Medicine

> "Evidence-based medicine is the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients. The practice of evidence-based medicine means integrating individual clinical expertise with the best available external clinical evidence from systematic research." ([SRRH96]).

Evidence-Based Medicine

[SRRH96] are at pains to point out that evidence-based medicine is not a "cook-book" approach but involves each doctor developing a personal methodology for seeking out and adopting best practice through a lifelong process of continuing medical education. However it is clear that organizational and resource policies can help or hinder this approach, for example collating databases and vouching for the quality of their contents, such as the *Cochrane Database of Systematic reviews (CDSR)*. It clearly makes sense to make it easier for doctors to access the information they need at the point at which they need it, although recognizing that their ability to extract, judge and utilize this information will depend on their degree of experience — hence it not being a "cook-book" approach.
As developed by [SRRH96], evidence-based medicine is a methodology for good decision-making practice. As such their book contains chapters on "How to ask clinical questions you can answer", "Searching for the best evidence", "Critically appraising the evidence", how the doctor can "apply this valid, important evidence in caring for your patient" and self-evaluation. Each chapter also has a section devoted to the teachers of such topics as well as to the learners of them. Given the huge volume of medical research available, its not surprising that the chapter on critically appraising the evidence is the longest in the book.

[SRRH96]

The contrast in paradigm between the old and the new paradigm is brought out clearly in [Evi92]. They characterise the evidence-based medicine paradigm as follows:

1. "Clinical experience and the development of clinical instincts (particularly with respect to diagnosis) are a crucial and necessary part of becoming a competent physician. Many aspects of clinical practice cannot, or will not, ever be adequately tested. Clinical experience and its lessons are particularly important in these situations. At the same time, systematic attempts to record observations in a reproducible and unbiased fashion markedly increase the confidence one can have in knowledge about patient prognosis, the value of diagnostic tests, and the efficacy of treatment. In the absence of systematic observation one must be cautious in the interpretation of information derived from clinical experience and intuition, for it may at times be misleading.

2. The study and understanding of basic mechanisms of disease are necessary but insufficient guides for clinical practice. The rationales for diagnosis and treatment, which follow from basic pathophysiologic principles, may in fact be incorrect, leading to inaccurate predictions about performance of diagnostic tests and the efficacy of treatments.

3. Understanding certain rules of evidence is necessary to correctly interpret literature on causation, prognosis, diagnostic tests, and treatment strategy."

([Evi92], page 2421)

The authors go on to describe how the internal medicine program at McMaster University has designed its programme to teach these principles (largely conforming to the rallying call of [Bul93]). They correct various misinterpretations of evidence-based medicine, e.g. that it ignores clinical experience and clinical intuition. They describe barriers to teaching the approach (e.g. "for many clinical questions high quality evidence is lacking"), and barriers to practicing the approach (e.g. "economic constraints and counterproductive incentives"). In support of the former issue, [PMS...97] suggest that, in relation to judgements of the risks of invasive cardiac procedures between generalists and specialists, one possible explanation is that "Disagreements about the risks of procedures may arise from the paucity of published data or from an oversupply of confusing, contradictory data." Among other issues of support, they point to the need for effective computer-based technology. The consequential increasing reliance on electronic forms of information storage and retrieval (including patient records), raises issues about the effective design of systems for use by doctors in their decision-making role (see e.g., [EFC97]).

3.4 The Role of Academic Knowledge

Perhaps the least explored of the new approaches concerns the issues of how theoretical knowledge is used in practical situations. The use of theoretical knowledge appears to be largely a tacit process, but it is unclear whether this lack of clarification is the result of researchers' neglect or some inherent property. There is some evidence that use of theoretical knowledge increases at the consultant level; but could the delay be shortened by teaching specifically aimed at linking theory with practice, as suggested by [EABW95]? There is also evidence that theoretical knowledge is more deeply "encapsulated" in experts and is only used when other methods fail (see e.g., [SB93]). For example, [vLMP...95] found that there was a fall-off in scores of general practitioners on a written knowledge test from a highpoint 5–10 years after certification, except in the area of chronic illness. Their study was a cross-sectional rather than a longitudinal one, and the results are hard to interpret reliably. It might be that the test favoured the more recently qualified in the kinds of knowledge it was investigating. It could also be that the knowledge tested was more deeply "encapsulated" in the more experienced GPs.

The term theoretical knowledge is used here to refer to propositional knowledge in the natural and social sciences which can be found in textbooks and the research literature. This constitutes the central core of what [Era97] describes as Type A knowledge — public knowledge which is (1) subject to quality control by editors, peer review and debate

and (2) given status by incorporation into educational programmes, examinations and qualifications. Type A knowledge includes propositions about skilled behaviour, but not the skills themselves. It is confined to knowing that and excludes knowing how. Professional knowledge, however, is defined as a form of Type B knowledge — personal knowledge which is categorised by the context and manner of its use, rather than its source or epistemological status. It is that knowledge which professionals bring to their practice that enables them to think and perform on-the-job. Thus Type B knowledge incorporates not only propositional knowledge (in the form in which it is used) but also procedural and process knowledge, tacit knowledge and experiential knowledge in episodic memory. Skills are treated as part of that knowledge, thus allowing representations of competence, capability or expertise in which the use of skills and propositional knowledge are closely integrated.

3.4.1 Transfer is a learning process

The technical terms most frequently associated with this process are "transfer of knowledge" and "application of theory". The implicit assumption is that one carries scientific knowledge across from an education context to a practice context, then simply applies it. Yet there is a large body of evidence (1) that knowledge may not be carried across by many professionals, because they do not recognise or appreciate its relevance to a particular case or problem, and (2) that they may not apply it because they do not know how to do so. Instead of seeing transfer as an event in which a person suddenly becomes able to apply knowledge acquired in one context to a second, different context, we have to see it as a learning process in which a person not only carries knowledge from one context to another but learns how to apply that knowledge in the new context. Knowledge is acquired in a particular context and remains situated in that context until it can be transformed and resituated in another context; and the extent of this further learning will depend both on the degree of difference between the two contexts and on that person's preparedness and prior experience of successful transfer. Even when transfer involves a sudden flash of insight, considerable learning effort may be needed to convert that insight into usable knowledge. However, programmes for professional formation seldom recognise the learning effort required for the transfer of knowledge. Support for transfer is rarely provided, even though trainee and novice professionals are ill-prepared to tackle it on their own.

Transfer of knowledge is most difficult when the contexts and modes of learning are very different. For example, moving from a university context to a professional practice context involves changing from an environment in which Type A knowledge is dominant to one in which Type B knowledge is dominant. Moreover, resituation of scientific knowledge will require not only transformation of that knowledge but also gaining sufficient understanding of the new context to know what kind of transformation is needed. This problem is further exacerbated by the likelihood that other areas of scientific knowledge and other forms of professional knowledge will also be needed for appropriate practical action; and these different areas and forms of knowledge will have to be integrated somehow. This is a more complex and challenging problem than learning to use scientific knowledge in well-defined situations in familiar academic contexts.

Attempting to deal with the problem of transfer has been one of the motivating forces behind the development of problem-based curricula.

3.4.2 Problem-based learning

Patel and her colleagues have explored the role of basic science in the curriculum of medical schools. They characterise two broad approaches — the "*science-oriented*, in which basic science is taught independently of the more clinical aspects of medicine" and the "*clinically oriented*, in which basic science is taught within some appropriate clinical context" ([EP89], page 53). They report a series of experiments (see e.g., [PGS88, PGF86]) in which subjects of varying medical expertise (from first year students to expert researchers) had to read and explain a description of a medical case.

They found that "the only group that shows reliance on basic science is the intermediate-novice group. Experts and advanced-novices use clinical-situation models, which can be elaborated to accommodate basic science; though the novices are less selective in their application of basic-science knowledge. Beginning-novices cannot use basic-science information." (page 105).

In their view basic science and clinical science are not related hierarchically, with basic science the more fundamental and offering explanations at a finer level of granularity. The two approaches are to be regarded as separate domains, providing different *kinds* of explanations and, as sciences, providing different *kinds* of generalization. In a similar, later study of medical students [PEK90] found that basic science was utilized more effectively as a framework within which to construct explanations than as a method for facilitating problem-solving.

A direct comparison of the science-oriented and clinically oriented approaches in two medical schools is provided by [PGN91]. They compared 54 students from McGill — following a conventional curriculum (CC) where the basic sciences are taught first in separate disciplines — with 54 students from McMaster — where there was an established problem-based approach (PBLC). The methodology was as above, namely examining the explanations of clinical cases, read by students in different years of study, to see how they integrated biomedical information. There were two conditions: in one, basic science material was provided *prior* to the clinical problem; in the other it was provided *after* the clinical problem.

They found that diagnostic accuracy increased with seniority of the student for both medical schools. In particular

> "The PBLC students had learned a systematic process of thinking that was explicitly taught. The predominance of backward reasoning, the systematic use of clinical information, and the tendency to formulate extensive elaborations are all consistent with the notion that the students were generating diagnostic explanations through the use of hypothetico-deductive reasoning. The fact that these patterns appeared with beginners and did not change with level of training supports the notion that they reflected reasoning strategies taught at the beginning level and reinforced throughout the curriculum.
>
> ...
>
> Since systematic reasoning is not taught in the conventional curriculum it was somewhat surprising to see a systematic method of reasoning emerge, characterised by a relative prevalence of forward reasoning and a tendency to refrain from extensive elaboration. The CC students also showed a pro-

> nounced tendency to explain the case on the basis of a single diagnosis rather than an extensive list of differential diagnoses." ([PGN91], page 387).

[PGN91] also point out that the PBLC students had a tendency to more elaborate explanations, but also to generate errors within those elaborations.

[AM93] cite the above study along with over a hundred others in a detailed meta-analysis of studies (1972–1992) of the effects of problem-based learning. Their overall conclusions are:

> "Compared to conventional instruction, PBL, as suggested by the findings, is more nurturing and enjoyable; PBL graduates perform as well, and sometimes better, on clinical examinations and faculty evaluations; and they are more likely to enter family medicine. Further, faculty tend to enjoy teaching using PBL. However, PBL students in a few instances scored lower on basic sciences examinations and viewed themselves as less well prepared in the basic sciences than were their conventionally trained counterparts. PBL graduates tended to engage in backward reasoning rather than the forward reasoning experts engage in, and there appeared to be gaps in their cognitive knowledge base that could affect practice outcomes." ([AM93], page 52)

Because only three of the studies examined by [AM93] dealt with performance assessment of graduates, they felt unable to draw firm conclusions but urged that further research was needed in this area.

In a larger and more carefully controlled study than that of [PGN91], [SMBH...96] compared 612 Dutch students diagnosing from 30 epidemiologically representative clinical case descriptions. The students were drawn from three Dutch medical schools — one teaching a conventional curriculum, another taking a problem-based approach and a third a *systems* approach that "integrates the biomedical and clinical sciences around major organ systems. Students engage in patient demonstrations and small-group training sessions in which knowledge previously acquired is applied to relevant clinical cases." (page 660). Unlike [PGN91], these students were not exposed to basic science material as part of the study. [SMBH...96] found that in all three schools, students produced monotonically better diagnostic performance by level within the school. They found a significant effect of curriculum type on diagnostic performance, as well as an interaction between curriculum type and year of study. A post-hoc analysis showed that students taught the conventional curriculum performed more poorly than the other two groups (the problem-based and the systems-based).

The authors surmise that the reason for the differences in outcome may be that:

> "the problem-based and the integrated curricula [i.e. the systems-based] both offer subject-matter to students in an integrated fashion and that students are encouraged to process the information in an active way through small-group discussion. Thus, subject-matter integration and active processing seem more important factors in attaining proficiency in diagnostic reasoning than the amount of self-directedness in the curriculum. ...In addition, seeing paper patients or their real-life equivalents seems to be important." ([SMBH...96], page 663)

Results conflicting with the above, but using ratings of clinical supervisors, were found in a study of 485 Australian graduates. "Graduates from the problem-based medical school were rated significantly better than their peers with respect to their interpersonal relations, 'reliability', and 'self-directed learning'. Interns from one of the two traditional New South Wales medical schools had significantly higher ratings on 'teaching', '*diagnostic skills*' and 'understanding of basic mechanisms'." ([RAP...95], page 225, our emphasis)

How far the different courses or the different methodologies used in the Dutch and the Australian study contribute to their conflicting result on diagnostic skills is difficult to assess.

3.5 Summary

Theories of expertise developed in different contexts using different research techniques may emphasise different aspects but do not greatly differ in their conclusions. Key features include the importance of case-based experience, the rapid retrieval of information from memory attributable to its superior organisation, the development of standard patterns of reasoning and problem-solving, quick recognition of which approach to use and when, awareness of bias and fallibility; and the ability to track down, evaluate and use evidence from research and case-specific data. Understanding the nature of expertise is important for self-monitoring one's use of heuristics and possible bias, sharing knowledge with others and supporting other people's learning. It is also critical for understanding the respective roles of clinical experience and evidence-based guidelines.

Chapter 4

Learning in Clinical Settings

Doctors learn to become competent and to make good judgements through the experience of treating patients with appropriate levels of guidance and supervision. They also learn by consulting and studying textbooks, reference books and journals. Then thirdly they participate in formally organised teaching and discussion. All three can be found in clinical settings; studying and teaching also occur in other settings — library, seminar, lecture room, and a range of workshops or laboratories. This chapter focuses only on clinical settings, recognising that there is a continuum from treating patients under the watchful eye of more senior doctors, through engaging in discussions about those patients, to participating in discussions about patients one has not treated or being taught about the diagnosis and management of a patient in a mini-seminar or lecture soon after visiting that patient. The use of literature informs both the treatment itself and preparations for presentations and further discussions about that treatment.

Three modes of progression can be usefully distinguished;

1. Extending the range of cases one is competent to handle;
2. Increasing the complexity of those cases; and
3. Increasing the level of one's responsibility for the management of each type of case: from peripheral onlooker, to the acquisition of evidence from and about the patient, to provisional decision-maker, to treatment under supervision, to virtually total responsibility.

A well-conceived training programme will take all three modes into account. This involves the selection of clinical settings, the selection of patients for attention by the trainee, the trainees' level of responsibility for each type of patient at a given stage in their training, and the nature and amount of supervision or guidance or teaching given. Another critical factor is continuity of care: over what period of time does the trainee have contact with and/or responsibility for various types of patient; and is this sufficient for the proper development of competence? [LM91] report improvements in medical care as residents gain longer periods of contact with patients.

Although more formal attention is given to opportunities for clinical experience when selecting clinical settings, consideration may, indeed should be given to the wider clinical

environment. The ethos and values of the workplace and the prominence of good role-models can have a major impact on the professional development of doctors. The role-model aspect is more formalised in the North American residency programmes through the evaluation of clinical teachers (see pp 56–58). A related issue concerning individual teachers as well as training programmes concerns the quality of supervision. While most research has focussed on the supervision of house officers (Section 4.1) and trainee surgeons (Section 4.2), its disappointing findings may apply equally to registrars and trainee physicians. Supervision is not only central to the role of a clinical educator, but also an important extension of patient-centred values to conduct beyond a supervisor's direct personal contact with patients.

Research into the outcomes of training programmes is sparse, although it can serve a range of useful purposes. One of these is establishing realistic standards of competence, so that none of the parties affected — the doctors, their patients or those who employ the doctors — misunderstands the level and range of their competence. The difficulties of achieving even this level of information are discussed in the section on assessment; but it is nevertheless useful to consider what use might be made of evidence already gathered by the formal assessment processes embodied in the Royal College examinations and the award of the Certificate of Completion of Specialist Training (CCST). Such evidence might contribute to the formative evaluation of training programmes, the planning of CME for newly appointed consultants and their duties when appointed, and help to validate (if possible) that most flexible of research instruments — the self-report. This is important, because the small numbers involved, in regional training programmes, and the diversity of experience within those programmes make it unlikely that controlled comparative research will yield significant results. Even if marked differences in outcome were to be observed, it would be difficult to attribute those differences unambiguously to particular programme features.

4.1 Evaluations of Postgraduate Basic Training in the UK

Pre-Registration House Officers (PRHOs) are technically postgraduates, although not yet registered as doctors, so they have been included. Studies of PRHO training emerging throughout the 1990s, used different methods and foci but nevertheless presented a consistent picture. Two major improvements have been a substantial reduction in PRHO hours of work and clear programmes of educational seminars; but progress in other areas has been slow. Four conclusions from [CD91]'s critical incident study still apply today: the need for all PRHOs to receive proper feedback on their performance, for access to someone senior to talk to, for training in working with dying patients and for greater use of the learning opportunities provided by their clinical experience.

A questionnaire survey by [GDA...93] of PRHO training in eight English regions combined questions about hours, conditions of work and inappropriate duties with questions about key aspects of competence and quality of supervision. While 91% and 61% reported receiving adequate guidance in basic and advanced cardiopulmonary resuscitation, only 65% (basic) and 10% (advanced) felt confident in performing it unsupervised. Only 41% reported adequate guidance in breaking bad news and 44% in pain control. Ratings on three aspects of supervision were 2.7 (briefing), 2.9 (problems) and 2.6 (career), all on the "dissatisfied" side of neutral on a 5 point scale.

[Wil93] studied the consultants' perspective on PRHO training, concluding that while they welcomed the concept of an educational supervisor "theoretically", many felt unable to take on this added responsibility. Teaching ward rounds were the standard method of instruction but most consultants estimated their total teaching time to be less than 30 minutes a week. Of the 33 respondents, only 8 had one to one teaching for their house officers; while 24 said they delegated some of that responsibility to other junior medical staff and to ward sisters. 25 said they would welcome feedback on their teaching and 18 would have liked training in educational methods and principles; and they were all concerned about not downgrading the apprenticeship system. Their problem did not appear to be their attitude towards teaching, but the pressure they felt from their lack of time to give to PRHOs when overwhelmed by service demands.

Two in-depth studies in particular localities by [DB91] and [E...97] focused respectively on the management and learning of PRHOs. [DB91] highlighted the range of expectations of PRHOs, the disparate and distributed nature of their supervision, their lack of clear goals and systematic training, and the absence of any coherent management of their work and training. [E...97] found that only a minority of 33 PRHOs in three hospitals received formal feedback from a consultant; and that even informal feedback was totally absent in important areas such as communication with patients. ([RSFC97], found a similar neglect of interactional skills during the intern year in New South Wales). They would have liked a wider range of clinical experience, especially in clinics and surgery (c.f. [TSO96]), and more bedside teaching. Teaching and learning in clinical areas was always on a case by case basis, not one respondent could recall reviewing a group of patients with a similar condition or engaging in what the Americans call a "chart review". Some received quite a lot of teaching on ward rounds, some virtually none. There was good practice in every hospital but also areas where teaching was neglected. There was no guarantee that official recommendations would be implemented and no quality assurance of training (c.f. [GSL97]). Organisational factors such as frequent changes in composition and distributed working locations prevent the formation of the kind of stable working teams with which the ideal form of apprenticeship is frequently associated.

An earlier survey by [GMK89] of the training of Senior House Officers (SHOs) reported similar levels of dissatisfaction. On a scale of 0–3, statements receiving agreement with means of over 2.0 included:

SHOs provide service more than receive training	2.6
SHO hours are not conducive to learning	2.6
SHOs do not think they are receiving postgraduate training	2.2
No one sees teaching as a defined part of their job	2.2

Only about 53% reported receiving most teaching from their consultants; and their views of the learning methods on which SHOs relied most differed from their consultant supervisors (see Table 4.1 below).

The authors note that the "apprenticeship" model, if unanalysed, "perpetuates the unhelpful confusion between training and service, to the extent that providing the service may become identified with receiving training". If consultants are to be given responsibilities as teachers or supervisors, the necessary time must be built into their routine schedule of work. Whereas SHOs seem to be regarded both as full-time doctors in the

Learning Method (24 less used methods not listed)	% included in top 5	
	Consultants	SHOs
Ward round with consultant **	84	55
Ward teaching with registrar/senior registrar	41	47
Formal teaching with consultant	23	19
Supervised practice	17	12
Unsupervised clinical practice **	19	44
Discussion with peers *	34	44
Library work or reference to journals *	25	37
Reading about cases seen *	33	50
Following up outcomes if cases seen	20	28
Presenting patients at small meetings *	24	12
Teaching in outpatients *	20	10
Teaching for Membership exams	20	13
(* difference of 10-20%, ** difference of over 20%)		

Table 4.1: Priority given to 12 methods of SHO learning by consultants and SHOs

hospital service and as students when off-duty, with the usual obligation to study in what little time is left to them.

[B...94]'s study of SHOs and registrars in four types of hospital concluded that variation in the content of training was caused as much by lack of structure as by type of hospital. Most consultant teaching occurred during business rounds, though the teaching element was rarely formally defined. There was great variety in the amount of ward-based teaching, depending on the propensity of individual consultants to teach. Similar ad hoc access to learning opportunities was found in outpatient clinics and theatres. The authors also noted that there was little attempt to cater for the differing educational needs of different grades of junior doctor. However [PWC...97] were able to show significant improvements in SHOs' rating of their post between 1992-3 and 1994-5: the proportion recommending the post to a friend as 'good' or 'excellent' rose from 45% to 57%, and ratings of their consultant supervisor rose from 48% 'good' and 'excellent' to 61%. There was also a large reduction in the frequency with which they felt forced to cope with a problem beyond their experience or competence: only 2% felt this at least once a day (previously 23%), 12% once or twice a week (17%) and 46% once or twice a month (28%). Nevertheless [PL99] reported that the new GMC guidance document **The Early Years** did not address the "real difficulty" of balancing service with teaching; and Calman (1999) made the same point about specialist registrars:

> "The balance between the education and service content of higher specialist trainees' work must shift if they are to be genuinely regarded as trainees". (page 31)

[BBM98]'s review of Basic Surgical Training in Yorkshire showed good consultant support for trainees, but deficient clinical experience was found in 32% of general surgical

and 70% of orthopaedic posts. This was due mainly to poor organisation of their clinical activity (for example 70% of orthopaedic SHOs never attended outpatient clinics), though it was not helped by the lack of PRHOs in orthopaedics. Only 8% of trainees had undergone an appraisal and "the concept of bleep-free periods for educational purposes was non-existent."

4.2 Learning Procedures

Wigton and his colleagues have published a series of papers on the learning of procedures on residency programmes in Internal Medicine in the U.S. ([Wig81b, Wig92, WBNT89, WBMN90]). Since these programmes are organised from (though not confined to) a single site he was able to survey the Programme Directors. His 1989 paper had two main purposes: (1) to find out from each Director which procedures **should** be mastered by **all** residents and which procedures **are** mastered by all their residents, and (2) to find out how many of each type of procedure were needed to attain and to maintain mastery. For several procedures the gap between **should** and **are** was 40% of all respondents; and many procedures commonly done in practice were mastered by all residents in less than half the programmes. The medium number of procedures required for mastery ranged from 5 for thoracentesis to 50 for chest roentgenogram interpretation. Once more a gap is revealed between performance on completion of training and the expected competence of a qualified doctor.

This raises questions about how the teaching of procedures is organised. One successful approach to this problem is described by [GBI91]. It involves (1) a requirement that procedures can only be conducted by a resident who has been formally credentialed for that particular procedure; (2) a phased timetable for acquiring the agreed list of necessary procedures; (3) regular feedback to residents on their current status and one-to-one consultation sessions for those who are falling behind. The proposal that "residents should be credentialed to perform, supervise and teach certain clinical tasks, particularly procedures, by performing those tasks under direct supervision" was approved by the Educational Policy Committee of the American College of Physicians ([Ame89]); and [GBI91]'s system made it more workable. The advantage of a credentialing system is that it requires competence to be demonstrated to those properly qualified to assess it. Not all aspects of medicine are suited to such a system, but it could be extended beyond procedures to include areas such as the administration of drugs ([BTR90]). Credentials for cardiopulmonary resuscitation skills ([MTWS91, WR93]) and others requiring regular practice should presumably be datestamped.

A further question relates to the context(s) where certain procedures can be most effectively taught and most validly assessed. Giving thought to this issue and arranging short term placements could be helpful. At a more advanced level, [BOSR96] report that learning nasendoscopy on awake patients in an ENT clinic is an excellent introduction to fibreoptic skills and transfers to nasotrachial intubation in anaesthetised patients.

4.3 Learning in the Operating Theatre

A Scottish study by [SLM89] on the technical training of general surgical trainees covered experience, supervision and indirectly competence. 47% of those in teaching hospitals reported that their operative workload was too little (23% in district general hospitals); 35% said they did too little operating alone and 58% too little supervised operating. More worrying was the response from 32% that they were "sometimes out of depth" when doing emergency operations alone in DGH (15% in teaching hospitals, 9% for elective surgery). As in the Gillard study, trainees' concerns relate both to the amount of **experience** (resuscitation or operating) and the degree of **supervision** (in ward or theatre), the two most central features of learning in clinical settings. Steele later participated in an audit of colorectal cancer surgery in three regions for the period 1990-4 ([ATS...99]) which revealed both low levels of supervision of registrars and SHOs and a much lower level of operating experience than that offered by an US fellowship programme for residents. This concern was backed up by a Scottish mortality study of avoidable deaths after colorectal surgery ([MNA98]), a finding which prompted [Col99] to suggest that surgical training deserved the priority attention given to waiting lists.

Concern about supervision in the operating theatre was raised again by a survey of unsupervised 'first time' procedures by [Wil97], who also raised the problem of excessive 'first time' surgery by young consultants and arrangements for the introduction of innovative procedures. In response [OS97] noted a cultural change towards defining a well trained surgeon in terms of the number of operations performed under supervision rather than the number of operations performed solo; while [GM97] noted improvements in training subsequent to Wilson's survey and the introduction of an assessment form to be sent directly from trainees to the supervisory body. Earlier [Ree93] had urged that careful evaluation would be needed as surgical training was changed from "a protracted apprenticeship, where endurance is eventually rewarded, to a shorter, more intense, residency style programme" (page 198s).

In the United States, [FWT93]'s audit of over 4000 cases showed a correlation between the absence of a supervisor and increased complication and mortality rates, but was not able to take case severity or complexity into account. Similar concerns about supervision have been raised about emergency departments; though [SBO...98]'s study of 3600 US patients revealed better compliance with guidelines when supervisors were present but no significant changes in patient satisfaction or reported problems with care. [For94] reviews how the urgent training problem associated with the introduction of Minimal Access Surgery was approached.

4.4 Vocational Training of General Practitioners

[HCG98]'s review of evidence from several countries of the effectiveness of vocational training, concluded that its introduction had improved the preparation of GPs, but comparisons between different models of vocational training were beyond the researchers' brief. The most searching study of programme outcomes was conducted in Holland by [GMHL...89]. This measured changes in the consultation skills and medical performance of General Practitioners during a one year programme of vocational training with a similar structure to that used in the UK, i.e. 4 days a week in the practice of a trainer and

1 day a week on a formal course. Both the pre-test and the post-test were based on 20 audio-taped consultations, rated on 26 variables by a panel of 4 assessors. Two groups of trainees were compared, a cohort of 32 from Nijmegen and a cohort of 31 from Gronigen. Significantly greater gains in medical performance resulted from the Nijmegen programme based on systematic skills development than on the Gronigen programme based on problem-based learning. Increases in consulting skills were similar for both groups in four areas but for "psychosocial performance", the Gronigen group maintained a high starting level without further improvement while the Nijmegen group improved from a lower starting level. Performance on knowledge tests also increased and was strongly correlated with both consultation skills and medical performance. Moreover, those with above average knowledge scores at the beginning made the greatest improvement in medical performance. Such improvements resulted from both an increase in obligatory performance and a decrease in superfluous performance, judged against protocols for the most common conditions. Neither programme achieved a desired level of performance across the full range of common conditions; and the vocational training programme in Holland was extended to 2 years.

The British system of vocational training became mandatory in 1982, but had been used in some regions for several years. However, mandation made it necessary to establish criteria for the approval and reapproval of trainers or to tighten up the existing local criteria. The benefit is documented by a before-after study in the West of Scotland ([KM92]), which showed an increase to 84% of trainees receiving a tutorial weekly (or most weeks), and a decrease to 4% of those rating their training as less than "fairly good". Most British research into the training of General Practitioners has focused on consultations in GP surgeries and the weekly tutorials given by trainers. The main aim of these tutorials is to improve the quality of consultations both from a medical perspective and in terms of communication skill. Since formal summative assessment at the end of the practice-training year and for membership examination of the Royal College of General Practitioners requires the submission of a video of several consultations, this aim has high priority for all concerned. The content of the tutorials typically comprises some combination of cases and issues raised by the trainee, a randomly chosen set of cases seen by the trainee in the previous week, a previously agreed medical topic, and discussion of the consultation process itself, often prompted by notes of a consultation at which both had been present or a video of a trainee consultation ([Gra98]). Urgent cases are usually discussed at the time but may be reviewed again at a tutorial. Trainers have all received training for their role; and will almost certainly make use of the construct of a patient-centred consultation.

[Sha90] and [ST91] have explored the parallel between trainer or trainee-centred tutorials and doctor or patient-centred consultations. However, both [Mar91] and [PC96] have shown that many experienced doctors are patient-centred in their consultations but trainer-centred in their tutorials. [Gra98] used both telephone interviews and videos of tutorials, accompanied by separate confidential comments from each participant, to study the nature and effectiveness of tutorials. This raised issues of style, focus, continuity and effective use of tutorial time, all highly pertinent to the training of trainers and meetings of local trainer networks. The system can be described as the regular, flexible and reflective use of two principles established as important for effective social skills training; regular systematic feedback and facilitation of self-assessment of both verbal and non-verbal communication by video recording. While patient-centredness is

usually advocated by the trainer and perceived as a desirable goal by the trainee, no particular approach to consultation is formally modelled and taught. General practice trainees are learning by doing as well as from tutoring, so it would be impossible to isolate the contribution of tutoring; nor would a control group study without training or without video be acceptable.

Not surprisingly, such a short training period cannot cover all conditions, so there have been many articles suggesting areas where GPs need more training. One of the best arguments has been in Palliative Care, where [LWLW96]'s survey found that only 15% of a sample of GPs had received a tutorial, and less than a third felt they had received adequate training in pain and symptom control; while 85% would have liked to have a placement in a hospice as part of their training. Another was in the area of rheumatic disorders where a large survey by [LPC95] showed that less than a half of GPs had received any special training, although such disorders account for 10% of an average general practice workload.

Two Australian articles are also relevant. A critical incident study of GPs six month into training and 12 to 18 months later ([SKD96]) demonstrated several major areas of positive change as well as some new or continuing areas of difficulty. The latter include gynaecology, pharmacotherapy and dermatology, the diagnosis of common complaints with uncommon presentations, the skill of managing difficult or angry patients, the organisation for the follow-up of patients with potentially severe disorders, and managing feelings of guilt over missed diagnoses or poor management. The second is [SKF...97]'s report on a remediation programme for trainees placed in the borderline zone at their final assessment. This provides additional intensive training and supervision; and having such a "middle pathway" clearly contributes to the maintenance of standards.

[ST98] report the use of portfolios by General Practice Registrars (GPRs) for a number of purposes: ongoing planning of a "curriculum" to meet the needs of the individual trainee, as a stimulus to reflection, as a mechanism to capture examples of significant learning in general practice, to reach the more difficult areas of a GPR's experience and to help feedback. They found that a portfolio was useful in establishing confidence during the first half of the year, provided that the relationship between trainer and trainee was reasonably good. However, it ceased to be useful when preparation for the MRCGP examination became the dominant feature of their planned learning.

4.5 Ambulatory Care Settings

This North American term includes both hospital outpatient clinics and private offices (general practice surgeries). These settings are now increasingly used in US residency training, especially in Family Medicine and Internal Medicine, because they "allow learners to:

1. care for patients seen primarily in outpatient settings - especially patients who have chronic illnesses;
2. observe the natural and treated progression of diseases through continuity of care;
3. practise health promotion and disease prevention strategies;

4. develop patient communication and negotiation skills; and

5. deal with social, financial and ethical aspects of medical care".

([Irb95], page 898)

We are especially indebted to Irby's comprehensive 1995 review of research into teaching and learning in **ambulatory care** settings "where 95% of doctor-patient encounters occur".

Several researchers ([Kra82, NSK...84, OSW93]) report better case mix, more continuity of care, more trainee responsibility for patients and more effective trainer-trainee interaction in private offices than clinics; and [OSW93] found that private office experience led to better performance on an oral examination of clinical competence. In general, however, there is great variation across ambulatory settings which affects both what and how much is learned. [Low86] lists the most critical factors as: the organisation of the clinic, interference from inpatient responsibilities; limited time, space and supervision; access to attending physicians for advice and feedback.

Whereas there is some discretion, at least in theory, over the pacing of visits to inpatients, clinics and office consultations are timetabled in a manner that determines the overall time spent with patients. A major concern, therefore, is the time available for interaction between trainer and trainee and how that time was used. There is a significant amount of evidence from American research but little from Europe. One indicator is the proportion of cases discussed, another is the duration of the teaching component of that discussion. The proportion of cases discussed varies greatly with location and trainee's experience. Thus [WGSR88] reported consultation rates in a family medicine clinic of 48% for year 1 residents, 28% for year 2 and 26% for year 3; faculty saw 20% of the patients of first year residents and 12–13% of patients of year 2/3 residents. The average duration for discussion was 8 minutes for year 1 residents and 6 minutes for year 3. Consultations were more likely to be resident initiated as they became more experienced ([MWM78]). ([GG93]) found a 7 minute average for interactions involving presentation and discussion but a 15 minute average when the interaction included a visit to the patient. [KLZ...89] found that 50% of interaction time was taken up by case presentation, 25% by clarifying questions and 25% by discussion. These discussions were largely focused on patient management, differential diagnosis and/or patient compliance ([LM91, WGSR88, HVKA93]); even in family medicine clinics, family and psychosocial issues were less frequently discussed ([MM93]).

Three recent American papers discuss approaches to using the limited time for interaction between trainer and resident as effectively as possible. [FSB...97] focus on strategies during clinic/office hours, such as (1) scheduling patients so that some are seen first by the resident then jointly, while others are seen by the trainer alone, (2) case presentations in the presence of the patient followed by a short burst of oral questioning by the trainer, and (3) using reflection on experience to develop short teaching scripts. [MI97] discuss the selection and use of teaching points, role modelling and preparation for and follow-up of doctor-patient encounters. [DDS...97] also discuss strategies which do not impinge on clinic time, such as learner preparation for seeing patients, various forms of case conference, and linked independent study.

Adaptations of morning reports for ambulatory contexts are discussed by [PCS89] and [MJ93]. Residents pass 1 page summary reports of all patients seen each day to their

trainer, who then selects cases for discussion at a teaching session the following morning. They found that a wider range of medical conditions was discussed in the ambulatory setting and that greater attention was given to the medical interview and to social issues, increasing residents' reflection on their outpatient experiences.

Feedback is also a critical issue. [Irb95]'s review reported that a majority of students found feedback being given during 3-6% of case discussions, the majority of it being positive. However, [GG93] found that supervisors changed their views about diagnosis and management quite frequently if they visited a patient after a resident's presentation and discussion, often judging them to be more seriously ill and becoming less confident about the resident's performance. Moreover, [CLLB86] showed that counselling combined with detailed feedback on patient satisfaction scores significantly improved the performance of residents with below average scores, when compared with a matched control group. [EPE95] conducted an ethnographic study of trainer's corrections of interns' errors, which found four different patterns of response, all of which they describe as soft "face-saving" strategies, e.g. asking subsequent questions that contain hints to lead the intern to the correct answer. They suggest that, while successful in reducing social distance between teacher and learner, these indirect responses may fail to promote accurate self-assessment by the interns who may not grasp the central message about what was wrong. The observed behaviour was more compatible with [Miz84]'s report of the social denial of mistakes by house staff by reference to medicine as an art with a "gray area", repression, or redefinition of mistakes as non-mistakes than with a learning climate which facilitates learning from mistakes (see also [WFML91]).

Failure to use ambulatory settings was one major criticism of [NWMM95]'s survey of urological training in Australia. Hospital-based training alone left trainees ill prepared for outpatient or "office" urology, and without experience of non-acute conditions such as urinary incontinence or infertility. Communication skills and critical appraisal of the literature were not well developed and feedback on progress was inadequate. The majority of respondents felt that mentors required specific training to facilitate feedback to trainees. Positive changes reported in the US include the introduction of a clinical-hospice/palliative medicine rotation for fellows training in haematology/oncology ([VGVRGW95]), and the development of a Housestaff Ambulatory Block for internal medicine residents. This block covered areas prioritised by an analysis of unmet training needs — dermatology, rheumatology, health promotion, women's health, behavioural medicine, geriatrics and sexually transmitted diseases ([LJR97]).

Another area of research concerns the role and performance of the clinical teacher. A useful starting point is the policy adopted at McMaster University for the evaluation of the teaching performance of physicians on their residency training programme in internal medicine ([GNW...93]). 14 domains of performance were identified, and trainees' rating forms provided evidence to the Director of the Clinical Teaching Unit in each hospital. These domains were:

1. Role model of conscientious care
2. Role model of compassionate care
3. Support for house staff
4. Role model of practice of evidence-based medicine

5. Teaching of evidence-based medicine
6. Teaching of clinical skills
7. Teaching of biophysiology
8. Appropriate delegation of responsibility
9. Provision of feedback
10. Openness to feedback
11. Punctuality
12. Role model of respectful, cooperative, productive interaction with health care team
13. Appropriate use of consultants
14. Teaching directed at all house staff, junior and senior.

The inclusion of biophysiology reflects the determination in this particular programme to emphasise basic science as well as evidence-based medicine for which it is renowned. The other categories map well onto [UBS94]'s four roles of effective clinical teachers in family medicine: physician role model, effective supervisor, dynamic teacher and supportive person. [Irb95] used these roles as a framework for his summary (below) of research on teachers in ambulatory settings, concluding that all 14 studies reviewed were congruent both with each other and with studies of inpatient teaching.

Effective Teachers in Ambulatory Settings

Physician Role Model: Excellent clinical teachers served as positive physician role models. Such physicians were characterised as being knowledgeable and clinically competent, having good rapport with patients, and being perceived by residents as good role models.

Effective Supervisor: In the role of supervisor, effective clinical teachers gave residents responsibilities for patient care, provided opportunities to do procedures, reviewed patients with residents, and involved the residents in patient care. Excellent supervisors and teachers provided direction and constructive feedback as well. Greater delegation of responsibility for patient care was significantly associated with higher overall ratings of teachers by medical students.

Dynamic Teacher: Excellent clinical teachers were interested in teaching, were enthusiastic, made an effort to teach, spent time with individual residents, were available, engaged in dialogue with residents, provided explanations, and answered questions. Characteristics identified in other studies include availability and accessibility, organisation, clarity and enthusiasm, asking questions and giving directions, and directing residents' learning.

Supportive Person: Excellent clinical teachers demonstrated support for the resident, were easy and fun to work with, were friendly, and were helpful and caring. Characteristics mentioned by other authors include establishing rapport with learners, demonstrating a positive attitude toward teaching and motivating learning, and creating an educational environment that facilitates learning.

[Irb95]

[UBS94] also found that first year residents preferred more didactic teaching and one-to-one discussion, whereas third year residents took more responsibility for their learning and wanted more information and explanation.

To conclude this section, we draw attention to the report by [BBIS98] on a policy conference convened to discuss research needs for redesigning education in ambulatory settings. This reveals how little is known about this topic, even in North America. The members of the conference gave the highest research priority to the outcomes of ambulatory education, followed by teaching strategies, faculty development and the influence of organizational culture.

4.6 Inpatient Settings

We were unable to find any more recent survey of SHO learning patterns in hospital settings than [GMK89], which we reported in Section 4.1, nor any comparable study of registrars. Nevertheless that survey reminds us that unsupervised clinical practice and discussion with peers are important sources of learning as well as planned (or unplanned) teaching by senior doctors. This needs to be kept in mind when reviewing the research which is (a) dominated by studies of teaching and teachers and (b) largely from North America. In Britain patient-based teaching takes place mainly on the wards, but also at weekly conferences/seminars where case presentations are made to a wider group of doctors than the resident's own firm (there are called 'grand rounds' in North America). Ward rounds are described as 'teaching rounds' or 'business rounds', but the distinction between them is not precise and most ward rounds are described as a mixture of the two. Some of the time on teaching rounds may be spent in spaces close to the ward but away from the patient, if they are available.

In North America there are four types of activity which correspond to ward rounds: attending rounds, patient management rounds, teaching rounds and morning reports. Attending rounds involve the residents and an "attending physician", the American term for a fully trained doctor with overall responsibility for the patient (the nearest American equivalent to a consultant). Patient management rounds are defined as patient care rounds without an attending physician. Teaching rounds are rounds formally dedicated to teaching, but in practice often combined with attending rounds. [SM86]'s survey of 123 residency programmes in medicine showed that both the content and the location of teaching and teaching/attending rounds were identical. According to the chief residents, 15% of teaching rounds were only at the bedside, 22% only in the conference room, and 63% used both locations. The trend towards greater use of conference rooms was viewed negatively (though in Britain where the balance is more ward-based, access to other space might be envied). Morning report is a distinctly separate activity in North America, whereas in Britain its function is assumed either by informal conversations between consultants and registrars or by a scheduled ward round. The former would be a purely "business" transaction, the latter might also be used for teaching. [SM86]'s review indicated that in about half or more teaching hospitals, about half a day 5 days a week was occupied by rounds and morning reports. Attending rounds were conducted 5 days a week and usually lasted an hour or more. Teaching rounds averaged 3 days a week and a length of 50–90 minutes. Most respondents considered the optimal length to be 2 to 3 hours. Residents assigned greater educational value to (1) morning reports (2)

teaching rounds, programme directors valued (1) grand rounds and (2) teaching rounds.

Differences of perspective were also found on a smaller scale study by [WA95] using Critical Incident Interviews to investigate factors facilitating or restraining learning by residents. Both residents and faculty emphasised the importance of interaction and reflection in discussion. Faculty also emphasised concrete experience (with some support from residents), technical rationality and faculty expertise; whereas residents stressed faculty involvement and commitment, consideration of multiple perspectives and personal relevance. Two major restraining factors for residents (with some faculty support) were insufficient time and opportunity to learn and low faculty involvement and commitment.

[RL97]'s paper on Evidence-Based morning reports describes an innovative development of the teaching round to incorporate searches for research evidence and feedback on outcomes. Each morning report session lasts an hour and is divided into four phases: reporting search results from the previous day, reporting on new admissions, case presentations of three new patients (selected by the residents because they provided the group with clinical challenges whose resolution required research evidence), and formulating questions for a literature search and report back next day. On Fridays the third phase is devoted to follow-up reports on all the cases presented earlier in the week. Evidence of the impact on patient outcomes is currently being collected, but meanwhile a major change in departmental climate is reported, particularly in the emphasis given to evidence-based medicine and patient preferences and values.

Another adaptation of traditional morning report practice is reported by [War95] who also incorporated literature reviews commissioned the previous day to encourage the paradigm of evidence-based medicine. His prime innovation, however, is the presentation of a "discharge case" selected by the attending residents to get a good case-mix each month. The case review is expected to cover the reasons for admission, decisions and progress made each day in hospital, discharge and follow-up plans and costing for the hospital visit. The rationale was the lack of attention given by traditional morning reports to patient outcomes, psychosocial issues, humanness of the care rendered, rehabilitation and cost. Both the innovation and the new balance it gave to the programme were positively evaluated by the residents. [War95] was partly responding to [WS93]'s follow-up study of morning reports. They found that 24% of medical cases discussed at morning reports did not have a final agreed diagnosis when discharged. A review 6 month later revealed that 39% were still uncertain, while 36% now had a confirmed diagnosis which differed from the best "morning report diagnosis". Moreover 24% of those cases that had reached a "final diagnosis" now had a "late final" diagnosis from that first proposal (17% changed while in hospital and 7% subsequently). Their conclusion was that "misconceptions of disease presentation and appropriate diagnostic evaluation are likely to occur if cases are never revisited."

[WAC...96] conducted a controlled trial with First Year residents on an Emergency Medicine rotation comparing bedside teaching alone with bedside teaching and written course materials. This showed learning gains on a written test but no greater gain for the group given the materials; and the satisfaction rate was higher for those without materials! However, the provision of on-line materials in emergency departments, designed for rapid consummation rather than in-depth study proved popular with House Officers in several British hospitals. In this case the main success criterion was HOs

using the facility and finding it helpful ([GM92]).

[GRHW97] reviewed research studies on the relationship between internal medicine housestaff training and patients' outcomes, examining in particular whether trainees' inexperience or their their workload affected affected patient care, the effects of the structure of the teaching service, and possible benefits of having an internal medicine training programme. They were unable to reach any valid conclusions and were rather wary of studies more than 10 years old.

The introduction of cognitive science perspectives (see Chapter 3) has led to a more sophisticated analysis of interactions on teaching rounds and case discussions. Thus Irby's detailed analysis of instructional thinking and decision-making by six highly rated but very different physicians conducting teaching rounds in a conference room following a resident-led ward round incorporates the concept of "teaching scripts". Though not claiming a representative sample, his framework of analysis is well matched to teaching in clinical settings and suggests many areas for further research.

Clinical Teaching Decisions by Six Physicians

Planning: informed by periodic learning needs assessments with their team, they set priorities in advance for the allocation of time. Half of them contacted the resident "on call" the night before for an up-to-date review of their patients. Decisions were made about which cases to highlight (for difficulty or for educational value) and which content to teach (team requested, teacher-determined, filling gaps in the team's knowledge). Five of the six read some literature and four prepared simple learning material (handouts, reprints, exercises, slides, etc.)

Diagnosing the Patient's Condition: the main difference from the diagnostic context described in Section 1 is the greater reliance on second hand information from other members of the team.

Diagnosing the Learner's Understanding: inferred from the presentation and a few key questions. One physicians commented that "the effectiveness of the presentation is directly proportional to how much I can remember of it", another that if one cannot follow the script, the thinking is probably disjointed.

Interactive Thinking: this involves continually monitoring learners' responses and adjusting to their needs, making mid-action decisions about educational opportunities, selecting from their repertoire of teaching scripts, etc.

Teaching: four of the six trainers organised their teaching around cases, but appeared to use cases to trigger "fixed scripts" for the topic, which they then improvised around using a "Socratic style of interchange". The key teaching points for a given case (the only protocol case used in the study) varied greatly between physicians.

Reflecting: all trainers reported later reflection on their teaching rounds, some about the patients, some about the house staff, some about both.

[Irb92]

[Hew91]'s report of an out-of-phase interaction between a resident and a trainer was based on their reflected comments, and suggests how material for constructive reflection on clinical teaching might easily be generated.

The eight principles of learning advocated by Irby's six physicians were: "actively in-

volve learners and ask lots of questions, capture attention and have fun, connect the case to broader concepts, go to the bedside, meet individual needs, be practical and relevant, be selective and realistic, and provide feedback and evaluation" ([Irb94b], page 336) to which he later added "model professional thinking and action" and "create a collaborative learning environment" ([Irb94a]). Their practice demonstrated consistent use of these principles. Moreover, they often used their knowledge in an encapsulated mode, making a precise teaching point but elaborating on the deeper structure of that knowledge, when needed or asked. [Irb94b] also advises that, while faculty development workshops are effective in developing general teaching skills and improving understanding of learners, "knowledge of case-based teaching is best learned in the context of departmental teaching, improvement and mentoring programmes where content-specific, case-based teaching scripts and strategies can be shared" (page 340). Irby's expert teachers said it took them six or seven years to discover a teaching style they felt comfortable with. With more local sharing of expertise that time might become a little shorter.

4.7 Summary

There have been several evaluations of postgraduate basic training programmes in the UK. Though there have been a few improvements, the overall impression is still negative. Many features of the educational policy seem to be appropriate, but they are not being implemented in many hospitals. There is insufficient supervision and feedback. Educational goals are subordinated to service demands. While many house officers receive good clinical teaching, a minority do not and assurance of educational quality is weak. Learning goals are only specified at a very general level, so there is little clarity about priorities, especially at the PRHO stage. Management of the educational process at local level is given little time and little authority.

Continuing concern has been expressed about the survey evidence on basic surgical training. Sometimes the problem is too little supervision of operations by house officers, sometimes the house officers get insufficient clinical experience. Operating under supervision is seen as the most critical feature of learning to be a surgeon and there is not enough of it. Some authors recommend the greater use of 'skills labs' and simulators. (See also Chapters 3 and 5).

The learning of procedures in medical posts has been criticised for being too haphazard: there is often little continuity of experience and guidance is often provided by doctors who are themselves not very experienced. Systematic attention like that found in many parts of North America (when it sometimes includes a credentialing system for procedures) has not been reported in the UK.

Research in Holland into the outcomes of their 1-year training programme for General Practitioner registrars led to the extension of the programme to 2 years. Not surprisingly research investigating particular areas of expertise has resulted in long lists of needs for GPs continuing medical education (see Chapter 6. In areas such as palliative care and psychiatry the argument seems particularly strong, in other areas one might look to other ways of distributing more specialist expertise within primary care organisations. GP training contrasts favourably with basic hospital training in its ability to provide tutorial support on a regular basis; and the commitment to quality improvement appears to be much greater.

Outside general practice, there is much more in-depth research in North America than in Britain. This has given particular attention in the last decade to learning in ambulatory care settings, a term which covers both family medicine and hospital clinics. More use is now being made of such settings in order to give doctors a broader experience of medicine, especially when significant aspects of care are being moved out of hospitals.

continued over page

Summary — continued

The key issues emerging from this research are:

- The time allocated for trainer-trainee interaction and making the best use of that time;
- Methods of providing feedback;
- Qualities of teachers rated as 'good' by trainees and their colleagues.

Methods for finding time for trainer-trainee interaction cover both time created within clinics by patient scheduling and the use of clinic experience for later case discussion and chart review. The papers in this area include many useful practical suggestions as well as evaluations of practice. In general, learning in ambulatory settings was found to give rise to discussion of a wider range of medical conditions, and greater attention to the medical interview and to social issues.
Issues relating to feedback included in the findings that:

- Supervisors changed their views quite often when they visited a patient after having heard the resident's report;
- Counselling combined with detailed feedback on patient satisfaction scores improved the performance of residents with below average scores;
- The effectiveness of feedback tended to be reduced by "face-saving" strategies which ameliorated or even disguised the central message.

Qualities of good teachers inferred from rating studies can be grouped under the headings of Physician Role Model, Effective Supervisor, Dynamic Teacher and Supportive Person.
Learning in inpatient settings is also researched in greater depth in North America. One gets the impression that American residents receive considerably more clinical teaching than their British counterparts, but there is no British data to enable a proper comparison. The variation in the amount of training received by British trainees is reported as considerable, raising issues of quality assurance and trainee entitlement. The two major constraining factors on learning by US residents were insufficient time and opportunity to learn, and low faculty involvement and commitment. Innovations receiving strong positive evaluations included adaptations of the Morning Report system to incorporate the teaching of Evidence-Based Medicine; and case reviews of patients whose diagnosis had changed while in hospital or within 6 months of leaving hospital.
Research is reported on instructional thinking and decision making by highly rated clinical teachers. This covers planning, diagnosing the patient's condition, diagnosing the learner's understanding, teaching and interactive thinking during discussions with learners, and reflecting.

Chapter 5

Learning in Non-Clinical Settings

The previous chapter spanned a range of learning activities — learning by doing, learning from text and computerised information bases, learning through discussion, and learning from presentations and mini-lectures. The criteria used for describing the learning as being in a clinical setting were (1) its direct relationship with the assessment and management of current patients and (2) its location within the concerns of the health care team responsible for those patients. This chapter covers a similar range of learning activities, but the mix is rather different because the focus is more on the acquisition of knowledge, skills and values needed for serving patients one will meet in the future than those which are one's current responsibility.

5.1 Teaching Conferences

Although the teaching conference (sometimes called a grand round) is similar in many respects to the teaching rounds described above, the cases discussed are more likely to be recent than current and the majority of those present will have had no responsibility for those particular patients. Indeed the aims of such conferences are not always very clear.

[Luc96], one of the few authors to address this issue, argues that if conferences are part of the formal postgraduate training programme their educational effectiveness has to be justified. His own department of plastic surgery has developed two forms of conference which are more consistent with principles of adult learning than the traditional variants. This involved distinguishing between (1) a **cognitive conference** based on prior study of assigned topics and readings in order to develop propositional knowledge and (2) an **intuitive conference** to facilitate "the use of judgement or intuitive reasoning in problem solving". The cognitive conference provides the focal point of a study programme in which faculty as well as trainees participate, one role of faculty members being to design questions on the reading appropriate for each level of trainee. The intuitive conference is based on four to six cases chosen from the week's theatre list. Each case is assigned to a senior resident and an independent consultant not involved with the patient are appointed as discussants. The consultant's role is to develop the presenter's reasoning through Socratic discussion. These conference formats have been developed over time by a series of critical reviews and modifications. The cases are chosen and assigned by

the director of the resident training programme with educational needs in mind, not left to the choice of individual presenters or their supervisors.

5.2 Journal Clubs and Self-Study

Another approach to the development of cognitive knowledge used in many programmes of postgraduate education is the Journal Club. Evidence on their effectiveness was recently reviewed by [Alg98], who noted that a major goal was to teach critical appraisal skills as well as establishing a lifelong learning habit. Clubs with high attendance and longevity are characterised by mandatory attendance, availability of food, and renewed importance by the programme director. Residents who are taught critical appraisal report paying more attention to the methods and are more sceptical of the conclusions, and have increased knowledge of clinical epidemiology and biostatistics but studies have failed to demonstrate that these residents read more, or read more critically.

A more individualised approach which aims to link self-study with current clinical experience is advocated by [HLKM96a, HLKM96b] in an AMEE Guide on **task-based learning**. The term is used rather generically to refer to a curriculum of tasks designed to take advantage of the core learning opportunities of a particular rotation. More ambitious still is the use of an educational contract ([MRA...94]), in which a resident agrees a complete 2-year schedule of rotations and educational activities, focussed on her or his unique needs and designed to meet both residency training programme objectives and individual resident-identified objectives.

5.3 Skills Teaching

The principle of finding the most appropriate trainers for each particular purpose is well illustrated by [DDVG97]. They noted weaknesses in house officers' communication skill and management of primary care type problems in emergency departments; so they developed a training course for them taught by General Practitioners. The programme involves weekly 1 hour small group sessions and a visit to General Practice. Evaluation is by feedback questionnaire, subjective tutor evaluation and SHOs were developing consultation and communication skills appropriate to A&E, getting support from and giving support to other colleagues (a reflection of the rather lonely and stressful job of an SHO in A&E), and learning about the management and treatment of primary care problems.

Another approach to inserting the teaching of clinical skills within a crowded postgraduate programme was developed in Canada by [CYGR...88]. Its purpose is to provide an ongoing review of physical examination which can serve as a top-up for first year residents. The four month course is taught in three parts each week:

1. The course director conducts a 1-hour session with 4 senior residents, starting with a 10 minute presentation of a relevant article by one of them, followed by a physical examination of a particular body system with the patient present throughout. There is no obligation to cover the whole system in a single session

2. All four senior residents teach the system to their clinical clerks

3. The original presentation is repeated to an audience of all the residents, who then divide into four groups and go to the wards, where each group is taught the physical examination by one of the senior residents.

Further details are in the article. The course has received a positive evaluation from all those involved and undergoes continuing improvement in the light of feedback.

Two skill-based courses in surgery for residents have also been reported. [LHG...92] describe a technical skills course for first year residents in general surgery, which runs for 3 to 4 hours a week for 8 weeks and includes 2 'dry lab' sessions (suturing and stapling) and 6 'wet lab' sessions. Residents showed significant improvements on a technical test and were better prepared for starting theatre work. The second describes a 3-day introductory course in laparoscopic surgery ([MDZ...95]), in which morning sessions pursued cognitive objectives using problem based learning and checklist construction with tutorial support and afternoon sessions were devoted to practice with pelvic trainers. As the technology of simulation develops further, the role of such skills courses is likely to expand ([DFA...98]).

5.4 Lectures

In many countries lectures in postgraduate education are linked to preparation for College or Board examinations which test both knowledge and its use in clinical situations. Their effectiveness depends not only on the quality and timing of the lectures themselves but also on concomitant teaching in clinical settings and time for independent study. Systems tend to be demand led (or demand resistant) but are not easily evaluated as there are few volunteers for a control group.

The role of lectures in promoting clinical competence is somewhat indirect. Their outcomes are normally measured by knowledge-based examinations, with the application of that knowledge being revealed mainly in clinical settings. However, unless accompanied by private study, knowledge gains from lectures can be very low ([Wig81a]). An unusual though pertinent example of a controlled study comparing "lectures only" with "lectures plus" interventions is provided by [S...93] in the field of ethics education. This showed that following lectures by case conferences with an ethicist in attendance did not increase the knowledge gain but significantly increased house officers' confidence in handling ethical issues. [SGFL92] discuss the impact of such ethics training on the treatment of patients with 'do not resuscitate' orders.

A later study by [SM97] of a two year medical ethics programme for house officers, comprising monthly sessions alternating ethics morning reports with "didactic conferences" describes consequent gains in both knowledge and confidence, as well as high ratings by 96% of the participants.

5.5 Other Approaches

[CRV...96] report the development, evaluation and further improvement of a management training strand for residents in general preventive medicine. The original version comprised special projects, coursework, committee participation, administrative and

program assignments, a one month administrative rotation and community rotations with administrative elements. This was well received but a survey of graduates also elicited several suggestions for improvement and indications of which aspects of the course could be strengthened. This led to further rotations with professional manager mentors and a series of management seminars, later expanded with a 1-semester course targeting selected management competences.

A different kind of learning environment is provided by day-release courses for General Practice Registrars ([Jen94]). This national system has significant local variation but is highly responsive to local demand. There is a great deal of groupwork and sharing of experience, as GP registrars are more isolated than those in hospitals. It also offers the opportunity to update on clinical areas neglected in their previous hospital-based training, to study consultation and counselling methods, and to address some management topics. The introduction of summative assessment will have affected these courses, but so far there is only anecdotal evidence of this.

5.6 Summary

Research into postgraduate teaching and learning in non-clinical settings mostly comprises evaluations of a wide variety of teaching innovations, rich in ideas but not necessarily generalisable. In particular we would draw attention to improving the learning benefits of departmental conferences, developments in self-directed learning (more prominent for CME), skill-based courses in surgery, the use of GPs to teach primary care to house officers in Accident and Emergency departments, a system for teaching clinical examination comprising both seminar and ward-based components, and confirmation that various types of "lectures plus" teaching are more effective than lectures only. Several departures from the standard lecture format have been positively evaluated, as have variations on case-based departmental seminars.

Chapter 6

Continuing Medical Education and Lifelong Learning

An Ontario Survey of physicians ([DDB...83]) distinguished between two types of CME activity: (1) those in which physicians were able to participate locally within the community of practice settings, and (2) those of a more formal nature, often requiring travel. Activities reported by over a third of the respondents are listed below:

Informal, local, community-based CME activity	
Reading journals	98.8%
Informal consultants	83.5%
Reading texts	76.0%
Attending rounds	72.9%
Using drug company materials	42.5%
Using AV materials	37.4%

Attendance at formal or distant CME programmes	
Scientific sessions	71.0%
Formal hospital events	52.1%
Meetings of local medical societies	44.6%
Medical school CME activities	43.4%
Speakers programmes organised by drug companies or other agencies	41.3%

Table 6.1: Types of CME activity: informal (top), formal (bottom).

[DLM94] describe this pattern as still typical ten years later, though a few newer methods were also beginning to feature by this time.

Meanwhile a structured interview study by [OAHH89] of General Practitioners in Wales revealed over 90% reading journals and medical papers and extensive purchase of books for personal use (76%) or the practice library (76%). However "reading medical literature" just failed to reach their top five educational activities. The percentages rating educational activities as very valuable (1 on a 1–5 scale) showed a strong preference for informal consultation and discussion:

Sources of Information	
Contact with partners, such as practice meetings and discussions	63%
Contact with patients	43%
Practice meetings with health visitors, social workers, district nurses	31%
Postgraduate meetings, courses and symposia at local hospitals	29%
Informal hospital input	25%

Table 6.2: Sources of information.

54% had carried out performance reviews within their practices and 46% held their own educational meetings (33% at least monthly). Their wish for more contact with other groups was demonstrated by 74% being in favour of non-medical members of the primary health care team being involved in CME for GPs, and by 70% favouring joint educational activities with hospital doctors.

Four years later a telephone survey of 111 GPs (response rate 63%) asked respondents to state the most important influence on the development of their practice of medicine ([DWW94]), leading to a very different kind of response (the total of 156% is due to dual responses being double counted):

Influences on development			
Education events	37%	Several factors	8%
Colleagues	29%	1990 contract	6%
Reading	27%	Consultant letters	5%
In-practice meetings	24%	Other matters	12%
Experience	8%		

Table 6.3: Influences on development.

The increased influence of educational events can be attributed to greater attendance following the introduction of a postgraduate education allowance (PGEA) into GP Contracts (contentious because it was funded by reducing seniority payments, leaving no net gain in income). However, we may be dealing with a bimodal distribution because, when later asked to identify any particular education event(s) that had changed the way they practice medicine, only 54% were able to respond positively.

Specific questions about five different aspects of change in their practice in the last 3–4 years yielded positive response rates as follows:

The first two aspects are linked to new contracts and financial arrangements: 73% attributed changes in health promotion to financial arrangements, though greater awareness (27%) probably affected performance (31% integrating health promotion with consultations and 10% the reverse). Changes in the doctor-patient relationship included an almost equal number of positive and negative responses, the latter usually attributed to deteriorating conditions for work, especially time pressures. The diagnosis and investigation responses largely concerned increases in investigations (ultrasound 18%, gastrocopy

Aspects of change	
Changes in practice organisation	90%
Changes in health promotion	89%
Changes in treatment (including drugs)	86%
Changes in diagnosis and investigation	65%
Changes in doctor-patient relationship	54%

Table 6.4: Aspects of change.

17%, blood investigation 14% were the most common) and referrals (24%), though some decreases were also reported. The main reasons given were improved accessibility (29%) and patient demand (24%). Only the changes in treatment were attributed mainly to CME, as shown below:

Sources of Information Leading to Changes in Treatment (n=95)			
Journals	67%	Cost/audit	26%
Educational meetings	43%	GP colleagues	21%
Pharmaceutical reps	30%	Patient pressure	6%
Local consultants	28%	Other sources	19%

Table 6.5: Sources of information leading to changes in treatment.

Specific treatments changes mentioned by 10 or more doctors included ACE inhibitors (54%), treatment of asthma (24%), anti-depressants (17%), antibiotics (17%) and treatment of GI disease (10%).

A more recent Welsh study by [AOR97] interviewed a random sample of 50 general practitioners and 50 consultants (response rate 77%) about specific changes they had made in the preceding year in four key areas of practice: management of a common clinical condition, prescribing, referral and use of investigations. Reasons for making these changes were then elicited and classified. Each group provided an average of 3.6 examples with GPs giving about 3.2 reasons for each change and consultants 2.8 reasons. The distribution of these reasons between eight most cited categories is given below for type of doctor and category of change.

GPs were twice as likely to mention cost factors; and consultants were twice as likely to mention changes in technology/tests. Patient-centred changes were more likely to be "patient led" for GPs and "patient need" related for consultants. Professional contacts were equally divided for GPs between consultants and other GPs, while 72% of consultants' contacts were with other consultants: a total of 14% of these contacts were with non-medical professionals. The breakdown of reasons within the "education" category was as follows:

[DTOH95]'s review of randomised controlled trials in CME found 99 such trials completed by the end of 1994, incorporating both primary and secondary interventions. A third of these trials included residents in the target population; three quarters took place in ambulatory settings — private offices, care centres or clinics. 62% showed an improvement in at least one major outcome related to either physician performance or

Type of Reason	Percentage of citations for each category					
	GPs	Cons	Manage	Prescribe	Refer	Investigate
Organisation	19	17	18	13	12	9
Education	14	21	24	20	5	16
Contact with professionals	14	12	9	13	14	16
Patient-centred	11	8	12	7	12	8
Technology/tests	6	14	4	0	4	28
Economic	11	6	6	22	2	3
Pharmacology/ pharm. companies	8	9	13	18	2	0
Clinical experience	9	6	7	5	12	9

Table 6.6: Citations by category.

Educational category	Number of reasons cited	
	GPs	Consultants
Scientific or medical journals	13	36
Medical newspapers	17	0
General press	5	3
Other/unspecified literature	4	21
Attendance at scientific meetings/conferences	1	20
Postgrad/clinical meetings, GP refresher courses	25	8
Supervision of trainees	1	0
Research	4	6
Audit	4	6
Disease management protocols/guidelines	6	2
TOTAL	80	102

Table 6.7: Reasons for citation within the "education" category.

health care. Given the low criterion and the likelihood that only the more carefully planned activities are likely to have been submitted to RCTs, this percentage is remarkably low. However, 79% of the interventions using three or more strategies showed positive outcomes.

[WvdWG98] review of research on implementing guidelines and innovations in general practice confirms the effectiveness of multi-faceted interventions; and also noted that "many ineffective interventions involved the dissemination of educational materials or the provision of a short education programme" (page 963). The likelihood of positive outcomes is also increased by management support or the "social influence" of well-respected colleagues a finding consistent with other literature on the dissemination of innovations ([Rog95b]). The cultural significance of hospital-based interventions being targeted at GPs and being rejected as inappropriately didactic also has to be considered ([ST96]). Our discussion will begin with evidence on the effectiveness of CME courses then continue with a further section on other types of intervention.

6.1 Effectiveness of CME Courses

[DTOH95] also concluded that short (1 day or less) CME events usually bring about little change. Although there are a few examples of very short courses focused on simple practical skills leading to positive outcomes — [ACJ91], for diabetic retinopathy (4 hours), and [DWHO96], for oroscopy (2 hours) — most reported successes are of courses longer than 1 day. Some use pre-test/post-test evaluations rather than control groups. But in the absence of plausible alternative explanations for changes in physicians' competence it is reasonable to attribute these outcomes to CME events. Otherwise the RCT criterion would exclude a significant amount of evidence on the appropriateness of educational methods and the length of CME interventions. Four recent studies with GPs help to illustrate these and other issues.

[TR96] report that simulated tissue can be an adequate substitute when teaching minor surgical procedures, even though it is not very realistic. Their 2 day course was taken by 6 groups of doctors, totalling 52 GPs and 92 general practice trainees. It produced both knowledge gains and high levels of confidence in most of the procedures performed in the course, with joint injection being rated by 36% as the most important aspect of the course. But a follow up evaluation after 9 months ([RT96]) showed that confidence in joint injections had fallen (with the exception of the shoulder joint), while confidence in other procedures had remained high. This could have been due to lack of practice, because only a third of the attending trainees were now in practices offering a minor surgery service. The authors suggest that the course was ill-timed for those trainees who could not practice the procedures soon afterwards. The difficulty of setting up control groups was also confirmed when it emerged that 40% of their control group (not randomised) reported having had some minor surgery training elsewhere. The programme itself was also **improved** in the light of feedback, so the intervention did not remain unchanged, another laudable practice which interferes with the aim of the research!

An Australian study ([GSFHR95]) evaluated a skin cancer training programme for GPs of comparable length (3 sessions of 3–4 hours). The first session was an illustrated lecture/discussion on the epidemiology of melanoma, different forms of skin cancer and management options. The second involved a visit to a melanoma unit to accompany a specialist surgeon in reviewing new patients and follow-up examinations; and to familiarise themselves with health promotion literature and screening protocols. The third was conducted in a private outpatient clinic and focused on both diagnoses and Surgical techniques for excising skin lesions. Both experimental and control groups already had high ratings for "adequacy of excision" (a needs assessment problem for the providers?); and both increased the number of excisions performed by a similar amount (possibly caused by the attention effect of the pre-test serving as a reminder). The percentage of patients diagnosed rose significantly for the intervention group; and the diagnosis of slides improved from 53% to 67%. But the accuracy of this diagnosis, as judged by pathology reports remained at 60%. The authors conclude that changes in knowledge resulted from the CME course but were not translated into changes in practice. Though this is arguable for diagnosis and excision, it certainly applies to the screening rates, which also remained unchanged. Perhaps the treatment agenda outshone the prevention agenda - a common problem, even with GPs?

[CDFM95] conducted a randomised control trial to asses the effects of different educa-

tional techniques on the cancer control skills of 57 physicians. Methods used included interactive small-group discussion, role playing, videotaped clinical encounters, lecture presentations and trigger tapes. Performance was measured by using unannounced standardised patients (see chapter 8) with hidden microphones to visit one year after the programme. Significantly higher performance was found for breast cancer risk-factor determination and smoking cessation counselling: these were the areas where the CME programme had used techniques that rehearsed or portrayed and discussed clinical activities.

Another 10 hour training programme concerned the assessment and management of depression ([GUTW98]). This programme, designed for the Defeat Depression Campaign, included both specially developed video material and course material for follow-on workshops. All five sessions had 1 hour of presentation, viewing and discussion, followed by 1 hour of role-play consultations which were videotaped and discussed. Assessment before and after the course comprised (1) a consultation with an actor role-playing a depressed patient, (2) a semi-structured interview and (3) a Depression Attitude Questionnaire. Although there were changes in interviewing style and the doctors gained in confidence, there were no changes in the two key measures of systematic assessment. The actor-patient ratings indicated improved explanations of depression and better regulation of its management. There was no change in the use of cognitive intervention, which the trainers had observed as causing confusion during the training. Therefore the package was revised to give greater emphasis to assessment and to reduce the time spent on cognitive intervention. This change could be interpreted as an implicit recognition that the course was too ambitious for the time available; but this was not overtly discussed, so it reads more like a "common sense" modification. None of the other evaluations reviewed discussed the feasibility of achieving all the objectives in the allocated time, nor whether certain more important objectives needed to be accorded greater priority. Indeed [DLM94] reported that "no studies were found comparing outcomes by varying the course duration".

Reports of courses in communication skills show greater awareness of the time dimension. [EKS...87] described a short programme of two 3 hour sessions on general practice consultations, without including consultation practice, as a refresher course aimed only at updating research knowledge and changing attitudes. Their reported outcome of an increase in patient satisfaction is best interpreted as converting competence into performance, a less glamorous but equally important role for CME. An important pointer to current practice among, for example, GP trainer groups was [VHD79]'s study of a group of experienced GPs who met regularly to view videotapes of each other's consultations. Independent judges looking at their tapes before and after this groupwork found improvements in more than a half of the rated skills, although these higher scores were obtained at the expense of longer consultations. In the US a more focused course of two 4 hour sessions concentrated on detecting and responding to patients' emotional distress ([RHK...95]). These sessions included practice with simulated patients and reviews of 5 minutes sequences from audiotapes. Two experimental groups and a control group were used, one experimental group being taught a set of eight skills based on a Rogerian model and the other eight skills based on a cognitive approach to problem definition (PD) developed by [Les85]. Audiotape evidence indicated changes in practice involving the incorporation of taught skills, and this was confirmed by visits by simulated patients. Real patients of the "trained" groups of physicians showed greater reduction in

emotional distress for as long as 6 months after their visits (post-training) although the duration of these visits had not changed. "Trained" physicians recognised more psychological problems, used more strategies to manage emotional problems in their patients, and showed greater clinical proficiency in the management of a simulated patient, than physicians in the control group. Moreover, patients in the PD group showed greater improvement than those in the Rogerian group. It should be noted, however, that for the even more taxing goal of developing the counselling skills of doctors and nurses in cancer care [MF88] have found that a minimum of 3 to 4 days is needed.

6.2 Other Intervention Strategies

Educational materials are another major strategy for CME, but results for interventions comprising materials on their own are not encouraging. [DTOH95] review reported positive outcomes in only 4 of 10 RCTs but there are a small number of positive examples of materials affecting prescribing practices ([HSH...86, MvKMP91]). A recent study by [MLH98] showed that providing 210 urologists with 67 monographs over a 2 year period produced a significant improvement in scores on a knowledge test, which correlated only with the number of monographs read and years post-training. But the improvement was modest and the authors doubted if it had any clinical significance. Given the limited time available for journal reading and studies reporting low levels of journal reading among some groups of doctors ([WGW...89, Gor84]), the focus is now shifting to the use of medical information systems as the first point of access to published knowledge.

The impact of practice guidelines has been relatively well researched. [GR93]'s review covered 59 evaluations that met their criteria for scientific rigour: these were not confined to randomised trials, because for guidelines in particular "there is a danger that treatment offered to the control patients will be contaminated by doctors' knowledge of the guidelines" (page 1317). They found that all but 4 studies detected significant improvements in the process of care after the introduction of guidelines. Moreover, 11 of the studies also assessed the outcomes of care, 9 of which reported significant improvements. [DLM94] appear to have reached a different conclusion when they state that "the evidence for their (guidelines) effectiveness on changing physicians on patient outcomes **by themselves** is weak" (page 254). However, [GR93]'s tables include a column headed **Intervention (dissemination or implementation)** which indicates that in almost every study the circulation of guidelines had been accompanied by concomitant activities such as reminders, feedback or conferences (see below). The conclusion to be drawn (and because of their different criteria their sample is much larger than that of [DLM94]) is that when a set of guidelines is considered sufficiently important for its impact to be formally evaluated, it will almost always be accompanied by other activities.

The notion that guidelines are sufficiently similar to be treated as a single type intervention is somewhat naive. Not only do guidelines differ greatly in their format and intent, but in their clinical credibility; and the contexts for which they are designed may vary from those where potential users are impatiently awaiting their arrival to those where attracting the receivers' attention may be difficult. [GDT...98]'s study of Dutch GPs' use of a series of official guidelines, developed by working parties of experienced practitioners and specialists, focused on the relative importance of various attributes

of the guidelines themselves. Ten guidelines and 47 recommendations from them were selected for the study; then a volunteer group of 62 GPs recorded their decisions after each consultation for which one of the 10 guidelines was applicable. Table 6.8 below shows the percentage of compliance with the guidelines when each selected attribute was present and when it was not present; and the "strength of influence" factor was calculated from this data.

Attributes	**% compliance when**		**Strength of Influence**
	present	**not present**	
Controversial, not compatible with other values	35	38	0.26
Vague and not specific	36	67	0.24
Described concretely and precisely	67	39	0.23
Demands changing existing routines	44	67	0.20
Based on scientific evidence	71	57	0.13
Consequences for management	50	65	0.13
Demands new knowledge and skills	54	65	0.10
Will provoke negative reactions in patients	47	63	0.10

Table 6.8: Influence of guideline attributes on compliance

Another variable which affects the use of guidelines is involvement in producing them. [CH95] describe an initiative involving 12 general practices in Northumberland to improve the care of patients with diabetes. This involved doctors and nurses learning together, discussion of six practice audits, remediation of identified knowledge deficits and the collaborative development of protocols. Evaluation was based on practice records for 1986 (prior to the initiative) and 1991 (two years after the setting of standards). They reported that:

> "More patients received general practice care only or shared care in 1991 than in 1986. There was a reduction in the use of oral hypoglycaemic agents among non-insulin dependent diabetic patients and more patients were maintained on diet alone. A greater proportion of patients were referred to dieticians, ophthalmologists and chiropodists in 1991 than 1986, and there was increased recording of, examination of, and identification of, diabetic complications. Little change was found in the mean values for clinical parameters between the two years."(page 149)

A study of 92 GP trainers in the North of England ([Nor92]) found that clinical standards for common childhood conditions improved prescribing and follow-up for those GPs who helped to set up standards for that particular condition, but not for those who were not involved in setting those standards: the experimental design was well chosen for the issue, because each sub-group of GPs was involved in setting standards for a different condition, thus demonstrating that involvement in standard setting only resulted in positive outcomes relating to those specific standards

A London-based study disseminated locally developed guidelines on asthma and diabetes through practice-based education ([FGH...95]). Two groups of 12 practices received

guidelines and educational visits for either asthma or diabetes, thus acting as a control for the other condition. The intervention involved three lunchtime sessions: (1) introduction to guidelines and discussion of how practice management could be developed into a practice protocol with an emphasis on patient recall for annual review, a stamp for reviewing patients was given as a prompt, and home monitoring was also discussed; (2) review of the practice's organisational decisions, session on clinical content of guidelines, technical demonstration; (3) audit data from patient notes and further review. Practices receiving diabetes guidelines improved recording on all seven variables, those receiving asthma guidelines improved only on review of inhaler technique and prescribing. When the "prompt" was used, there was significant improvement in the recording of both conditions.

This London study involved both an educational intervention (educational visits) and an administrative intervention (the prompt), as recommended by [GR93]. Analysis of [GR93]'s tables reveals a predominance of administrative intervention with preventive care guidelines (63% reminders and 26% changes in patient records) and a slight predominance of educational intervention with clinical care guidelines (33% feedback, 29% conferences or seminars, 21% reminders, 12% changes in records). Based on their analysis of the effect size, the authors suggest that circulation of guidelines should be accompanied by a locally-based educational intervention and patient-specific reminders at the time of consultation. [DTOH95]'s review found the outcome evidence to be particularly good for reminders (significant effects in 22 out of 26 RCTs). Given our earlier discussion (Chapter 2) about the relationship between competence and performance, it should be noted that the educational intervention is normally aimed at developing competence by discussion of rationales, evidence and problematic or typical cases, while the administrative intervention is aimed at converting that competence into improved performance.

[DLM94] describe educational interventions involving visits to practices as "academic detailing", and report the use of pharmacists and nurses in this role as well as physicians. When the overt reason for such visits is administrative (though these will almost certainly be accompanied by educationally useful discussion) the term "facilitator" is often used. For example, the use of **facilitators** who visit practices to talk to practice nurses and improve office systems has proved effective in improving health promotion activities in Oxford for monitoring risks of arterial disease and in the US for early detection and prevention of cancer. Both studies involved control groups. In Oxford there were dramatic increases in weight and blood pressure measurement and identification of smokers ([FFG87]); and a three year follow up confirmed continuing attention to identified hypotensive patients whose number fell from 16% to 8%, but much less impact on smoking and weight ([MMF...89]). However, [ES97]'s systematic review of interventions for preventing coronary heart disease concluded that multiple interventions only have a significant effect on patient outcomes for high risk hypertensive populations. The New Hampshire study incorporated two interventions — a facilitated visit and an educational event for the physicians. There were significant improvements in mammography, breast examination, stool occult blood, 'reduce fat' recommendations, and 'quit advice' to smokers; and the facilitator intervention was both more successful and sufficient on its own ([DOK...92]). Similarly, a control study in Scotland by [BNC...95] demonstrated a significant impact for an audit facilitator on 12 representative general practices' diagnosis and treatment of childhood asthma.

An important aspect of many of these combined interventions is that they are trying to ensure not just competence but consistently good performance by addressing primary health care systems as well as individuals; and also by treating them as multi-professional contexts. Their goal is to expand **individual competence** into **team performance**.

[DLM94] conclude their review of the effectiveness of CME interventions by emphasising the intensity and complexity of interventions with positive outcomes and the multi-faceted nature of the change process. Learning in formal contexts requires (1) problem-based approaches to relevant clinical issues using authentic visual material, (2) formats for facilitating transfer such as small-group case discussions, peer review exercises in clinical settings, and especially role-playing or practice-rehearsal strategies which provide an opportunity to practice new skills and receive feedback, (3) incorporation of practice-useful devices to enable, remind or reinforce clinical physician behaviours; and (4) long term follow-up feedback on practice performance. Even from a purely technical perspective, which regards needs assessment and the design of appropriate CME provision as unproblematic, much CME practice emerges as based on wishful thinking by both providers and participants.

6.3 Needs Assessment and Audit

[DTOH95] identified five levels of needs analysis in the 99 trials of CME interventions (160) they reviewed; and these were associated with varying percentages with a positive outcome.

Level of needs analysis reported	**Number of interventions**	**% with positive outcome**
No clinical need reported	12	42
Identification of general clinical area requiring change, with clinical care references.	34	53
Based on nationally approved guidelines	41	61
Consensus agreement among local health practitioners	45	58
Gap analysis or targeted barriers to change	28	89

Table 6.9: Levels of needs analysis.

These findings suggest that needs analysis and clear targeting are important requisites for effective formal CME interventions; and hence that links with audit might be advantageous. But this does not imply that audit on its own will be effective. Indeed [DTOH95]'s review concluded that even audit with feedback was a relatively ineffective form of intervention (positive outcomes in 10 out of 24 RCTs); though it was more likely to be successful when the feedback was delivered in the form of a chart review.

The primary purpose of audit is quality assurance and improvement and it is normally applied at the team, department or site level. [DLG...91] describe it as a multistage process involving six steps:

1. Selecting a topic;

2. Establishing target standards or criteria against which a level of performance can be measured;
3. Observing practice by collecting, analysing, and presenting data;
4. Comparing performance with targets;
5. Implementing required changes through discussions, written policies or other mechanisms; and
6. Repeating the review to check that changes have been implemented and that quality of care has been enhanced.

In principle the process is iterative, and the cycle can be entered at any point. [DLG...91]'s "audit of audits" suggested that one reason for their lack of impact maybe the pressure to demonstrate to the authorities that audit is being done. They found that Stages 2 and 5 were frequently omitted from general practice audits.

An interesting example from the hospital context is a study of the Royal College of Radiologists' guidelines on referral practice ([Roy92]). Their Working Party piloted these guidelines in six centres and found them acceptable to local doctors, who also agreed to monitoring and review of their practice in relation to these standards by a local committee of peers and colleagues. However, most of the subsequent reduction in referral rates did not come from firms with a "high referral" record; and local committees were only prepared to intervene to encourage compliance by "high referral" firms in a quarter of the cases. Stages 1–4 of the audit cycle were implemented but in 74% of instances of apparently resistant high referral there was no Stage 5.

The most common method used in Stage 3 is a systematic review of medical records. Sometimes it is the only method used, giving rise to occasional confusion between the North American term "chart review" and the audit process in general. Chart reviews are used in teaching (See Chapter 4) as well as for audit; and their limitations are discussed by [TD85] who advise that they should rarely be used as the only source of evidence.

[LH94] review the uses of chart review under the headings of self-assessment, assessing physician competence (see Chapter 8), community based chart audit using comparative data and assessing the adoption of guidelines. They then identify three ways in which chart review can be used in conjunction with CME:

1. results of deficiencies from the review of physicians' charts can be used by those physicians to plan a group-oriented CME activity such as a lecture or workshop;
2. results of screening or peer review of charts can be used to provide individualised feedback and consultation to a physician; or
3. results of a variety of different types of chart review (particularly in hospitals) can be communicated to physicians responsible for designing CME activities. These CME planners integrate this diverse information into the selection of CME topics and the planning of CME activities.

The first method is the least used for two reasons. Chart reviews are rarely undertaken with formal CME activities as an expected outcome; and where deficiencies in physician performance are identified, a formal CME programme is only rarely the preferred

remediation virtually none of the 185 medical audits reviewed by [SBM81] were judged to need a CME response.

The second method, a form of audit with feedback (see above), is being increasingly used. However, [MBO91]'s review of 36 studies of the use of statistical information from audit or practice reviews suggest that it is most likely to affect practice when the recipients have already agreed to review that practice. Private mediation of information about performance is frequently reported or advocated, but not researched. In principle, chart review is also a good point to enter the audit cycle, but without a strong disposition to follow up the results the cycle never properly gets under way. Not surprisingly truncated versions of 'audit' undermine the whole concept.

The third method is also being used more often, but it depends on CME being well funded and organised. [CJ99]'s review of CME in general practice found 18 evaluations of audits with educational interventions, of which 17 showed a positive influence on doctor behaviour, but only one included data showing the behaviour change was sustained ([Pri98]). [Pri98]'s study is also of interest because the practice review was triggered not by a conventional audit of cohorts of patients but by a significant event focusing attention on a problem already recognised but not until then being given such high priority.

6.4 Self Directed Learning

Having explored the limits of formally organised CME interventions, we return to the normal ongoing pattern of mainly informal learning described at the beginning of this chapter. [JJM...94] distinguish three forms of physicians self-directed learning (SDL):

1. informal, ongoing, habitual activities directed to the maintenance of competence;
2. semi-structured learning experiences that typically have their basis in immediate patient problems; and
3. formal, intentional, planned activities.

Informal self-directed learning is considered by many doctors to be part of their daily routine. They involve journal reading, ad hoc conversations, interactions with drug or equipment company representatives, possible attendance at regular events like departmental or practice conferences. Their purpose is perhaps best construed as self-initiated scanning of the doctor's practice environment with no specific outcomes in mind. **Semi-structured SDL** is linked to immediate patient problems. Its focus is decision-oriented, finding the best way to manage a patient; and learning is incidental to that prime purpose. It may take the form of consultation with immediate colleagues or experts with whom they have some contact, or of literature searches and reading. This type of learning is closely tied to individual patients and their progress and therefore cannot be planned. **Formal self-directed learning** corresponds to what [Tou71] called a "learning project", in which there is a clear intention to learn about a specific problem or issue. There is a fairly clear sense of what needs to be learned, though detailed outcomes may be emergent rather than pre-planned ([GMS94]).

The first large-scale investigation to explore the reasons for physicians making changes in their practice and the learning entailed was [FMP89]'s interview study of 340 physicians in the US and Canada. They found that, although some reasons for change were personal or social, most changes were driven by the desire to be more competent in the delivery of healthcare to patients. Important considerations early on were: (1) developing an image of the outcome of changing, because if this was clear the process of change was more rapid and efficient; and (2) assessing what capability they needed for making the change, which partly depended on the level of excellence they hoped to reach. Planned efforts to learn were commoner when the capability gap was fairly large, and the nature of that gap determined how they decided to learn. Their main sources for learning were colleagues, reading and CME programmes; and usually they used all three sources in various combinations ([FDW94]) The implications for CME are the need: (1) to create opportunities for social interaction so that doctors can develop and sustain networks of contacts; (2) to develop medical informatics to facilitate literature searches: and (3) to ensure that CME events provide what self-directed learners need, as and when they need it ([HL92, ASST93, LBB...90]).

[SKRM97]'s survey of 118 doctors in the Dakotas and Minneapolis found that two thirds of their "learning episodes" were patient-centred and one third were concerned with gaining new skills and knowledge, thus distinguishing between learning triggered by the problems of a particular patient and learning to extend one's expertise for the benefit of future patients. The former fits Jennett's description of semi-structured SDL, but only the latter could be described as a learning project for which some planned learning, possibly including formal CME, might be needed. Both depend on the doctor's prior awareness of new (to them) areas of knowledge through ongoing informal learning by scanning the practice environment and familiar knowledge sources.

Earlier [GPW82] and [PC89] had outlined three stages in physician learning: (1) doctors deciding whether to take on a learning task to address a problem, (2) learning the skill and knowledge anticipated to resolve the problem, and (3) gaining experience in using what was learned. These stages were confirmed by [Slo99] using in-depth interviews with 32 doctors; though he found it necessary to add a further stage to take into account the requisite prior awareness acquired by scanning the practice environment ([JJM...94]'s informal SDL), and to allow for possible termination of a learning episode. This four stage theory of doctors' self-directed learning is summarised in Table 6.10 below.

Problem Type **Learning Stage**	**Specific (e.g., addressing a particular patient's need**	**Bodies of Skill and knowledge**
Stage 0: Scanning for problems and other interesting things	The doctor is aware that problems are "out there". The doctor is alert for problems which she might need to solve and, when potential problems are encountered, she moves on to the next stage.	
Stage 1: Deciding whether the potential problem encountered is worth pursuing.	The doctor senses a need for immediate action and decides on the spot whether to take on the problem; alternatively, the doctor reads a bit, talks with others and nevertheless decides quickly	The doctor feels uneasy ("Maybe I should review that stuff ...") and asks: Is this really a problem? Is there likely a solution to the problem? Are resources available so I can do the required learning? Am I prepared to make the changes in my practice required by the learning I do?
Stage 2: Learning what is needed to address the problem.	Learning involves reading (typically journals, less often texts) and talking with others who offer suggestions.	Learning involves comprehensive reading and taking available and appropriate courses.
Stage 3: Gaining experience using what's been learned.	Learning at this point means trying the problem solutions on the problem in question and seeing what happens.	Learning means trying the new skills and knowledge in a range of settings and gaining experience as a result. It also means reading again though now its purpose is not to learn the new thing but to see what kinds of experiences others have had with it.

Table 6.10: Slotnick's 4-stage theory of physicians' self-directed learning episodes

An important aspect of the model is that the nature of the learning is significantly different for each type of episode. A more elaborate version of the model (Table 6.11) examines each stage in further detail, addressing five aspects in turn: the doctor's overall **goal**, the discrepancy between that goal and the current situation, the **learning resources** likely to be used, the nature of the doctor's **reflective thinking** and the **criteria for completing** the episode. The model illustrates the complexity of both patient-centred and knowledge/skill centred episodes, and the character and purpose of the thinking at each stage.

Stage Attribute	**Stage 0 "Scanning"**	**Stage 1**	**Stage 2**	**Stage 3**
Goal	Identify potential problems to consider during next stage AND note issues potentially useful at some later point.	Decide whether the doctor should learn what is necessary to resolve the precipitating problem.	Learn the knowledge and skill necessary to begin resolving the precipitating problem.	Apply and become comfortable with what has been learned in resolving the precipitating problem.
Discrepancy	Doctor needs problems whose solutions will satisfy Maslowian needs.	The doctor lacks sufficient information to decide whether to pursue the problem's solution.	The doctor lacks the skills and knowledge necessary to begin resolving the precipitating problem.	The doctor lacks experience and/or confidence in what he is doing.
Learning Resources	All aspects of practice and daily life.	**Specific Problem:** The clinical situation, reading, discussion with other doctors. **New Body of Skill and Knowledge:** Reading, conversations, information at meetings.	For **specific problems**, primary sources are reading and consultation; for **new skill and knowledge**, they are reading, consultation and courses.	Primary sources are those used already and experience using the skills and knowledge learned. Doctors also seek others' experiences in similar situations.
Reflection	Focus on probable fit with doctor's life generally, practice in particular. Relationship of problems, information, and issues to practice is a central feature.	**Specific Problem:** Focus on patient, available reading and consultation; context varies from clinical and immediate to consultative and deliberate; purpose is to address the questions under *discrepancy*. **New Body of Skill and Knowledge:** Focus on information on skills and knowledge to be learned; context is consultative, deliberative, and occur anywhere; purpose is to answer *discrepancy* questions above	**Specific Problem:** Focus is knowledge and skill needed to resolve the problem; context is typically clinical and immediate; purpose is to learn procedures for addressing the problem. **New Body of Skill and Knowledge:** Focus is reading, context is deliberative and/or hands-on; purpose is to gain sufficient knowledge and skill to begin using the new learnings in resolving the precipitating problem.	*Post mortems* of what happened occurred in both specific problem and general body of skill and knowledge situations. The focus was on the doctor's experience as well as prior knowledge and experiences, the context is deliberative and may or may not be at the site of the action, and the purpose(s) to evaluate and gain experience. The experiences of others (both personal and published) is reflected upon as well.
Criteria for Completion	Problem, issue, or information seems interesting or important enough to be considered further.	Answers realized to: is there really a problem? Is there likely a solution to the problem? Are resources available for the physician to do the learn what is required to solve the problem? And is it practical for the doctor to do the learning?	Situation-specific indications: The problem requires action, resources were exhausted, others (e.g. instructors) told the physician it was time, there was nothing more to study; doctor specific indications: An acceptable plan existed, the doctor felt ready; the doctor was clear on what was to happen next; the doctor felt there was no value in additional learning.	All criteria were situation-specific in the sense that the doctor gained enough experience to be confident with the new learnings. This was evidenced by the doctor's attention shifting to other issues. The stage could also end because the precipitating problem resolved and the doctor lacked further interest.

Table 6.11: Attributes at each stage in a learning episode.

A less systematic picture emerges from a British study by [ARJ96] of general practitioners' reasons for changing their prescribing behaviour. This identifies a preliminary awareness of new possibilities from reading and brief discussions with other doctors, which may lead to action if it matches the doctor's preconceptions, comes from a highly credible source or is triggered by encounters with other doctors' prescribing practices or other critical incidents. But it also reports a rather precarious process of trying out new drugs, which appears to be highly dependent on the responses of a small number of early 'pilots'.

[Slo99]'s framework does not give early experience with a change quite such a precarious feel, but it was derived from a different population of doctors in a different context with greater exposure to evidence-based medicine; and was not specifically focused on prescribing. Moreover it offers a realistic and comprehensible approach to lifelong learning which could be discussed in detail with postgraduate trainees. They will need more than exhortation to help turn good intentions into workable patterns of Continuing Professional Development.

6.5 Continuing Professional Development

Most professions have now adopted the term Continuing Professional Development (CPD) as encompassing a wider range of learning experiences than those associated with the term Continuing Education, which still carries associations of more formal, provider-initiated, educational activities. Thus CPD comprises informal learning as well as formal learning, learning on-the-job as well as learning off-the-job, the full range of learning activities described in Table 6.11 above. [EMC98] suggest that CPD be defined as "all the further learning which contributes to how a qualified professional thinks and acts at work". They list its main purposes as:

- **Career development**
  This includes becoming more specialised; taking on wider professional responsibilities; directing research projects; becoming a clinical teacher, supervisor or trainee; becoming the management of a clinical department, academic unit or research unit.
- **Improving and maintaining the quality of practice**
  This includes both personal practice and contributing to health care teams and organisations.
- **Expanding one's domain of competence**
  Simply sustaining one's ability to discharge one's clinical capabilities in a manner which reflect the growth of knowledge and research evidence is a major challenge. Career development and role extension create further demands.
- **Facilitating changes in practice**
  The introduction of new approaches, techniques or modes of organisation.
- **Quality assurance for clients and the public** the expectations of professionals today include making explicit their previously implicit social contact with the public. Health care organisations are expected to have robust quality assurance

policies and to demonstrate their effectiveness. This includes ensuring the competence of all their members.

While some of these goals might properly be regarded as only the concern of individual professionals, most of them also affect the performance of health care organisations and the relationship between doctors and the public. The central problem of CPD policy is the reconciliation of four incontrovertible factors:

1. It is individuals who learn: their motivation, access to and use of learning opportunities and time for learning are vital.
2. Nevertheless social expectations affect their learning, learning is triggered by social events, especially encounters with patients' problems; much learning is from other people in a range of social contexts.
3. Health care organisations are both necessary for the provision of multi-disciplinary health care teams and support services, and legally liable for the quality of care
4. There is always strong public concern about the quality of health care, which therefore remains high on the political agenda.

The system adopted by many professional organisations of requiring an annual minimum quota of approved CPD activity has been used for General Practitioners, with the additional complication of linkage to a Postgraduate Education Allowance (PGEA). However, the PGEA scheme has been criticised for similar reasons to quota schemes in other professions:

- It focuses on attendance rather than learning.
- Events are chosen for convenience and minimum cost, often at the last minute ([MC97]).
- Choices are often based on areas of interest, rather than areas of possible weakness.
- Quota schemes tend to be inflexible, because activities other than CME events rely on self-report and cannot easily be confirmed.

Although such schemes have helped to build networks of professional contacts (important for informal learning) and to initiate a cultural change towards regarding lifelong learning as an integral part of all professional work, they have lacked a direct link to needs analyses based on professional performance. Hence they cannot be regarded as meeting all the demands for quality assurance. The current Australian Quality Assurance and Continuing Education Programme for general practitioners is of particular interest ([Sal97]). This is based on a 3 year quota of credit points, upon which continuing registration depends. In addition to formal CME and informal learning, there is a requirement for practice assessment activities based upon an audit cycle.

An alternative to the quota system is the personal development plan (PDP). The arguments for PDPs are that they are needs-based rather than provision-based and impose no restrictions on the type of CPD activity. However, they also lack verification of

outcomes and assurance that priority needs will be addressed, unless they are tightly coupled to audit or performance review. From another perspective, the PDP approach is often criticised for being too individually based. [OAHH89], for example, found that 58% of practitioners thought that self-learning activities on their own were an inadequate strategy.

One response to this problem of isolation has been to introduce a mentoring system. For example, the three-part strategy piloted by [CMHF97] for GPs in Sheffield comprised: the development of a personal education plan; creating a portfolio to document its progress and gain accreditation for the learning; and mutual support through a co-mentoring group, initially facilitated by a CME tutor. Other initiatives involving GPs' use of personal plans with support from a GP tutor or CME advisor are reported by [Bah98] and [VF98]. An East Anglian scheme which offered GPs a choice of an experienced GP mentor or co-tutoring in which pairs of GPs support each other's learning has been evaluated by [HB98]. Both systems were positively received, and 52% of the contracting group used PDPs compared with 27% of the mentored group: an additional finding was a significant reduction in stress levels in both groups.

Another response has been to introduce activities for groups of physicians who work together and for multi-professional health care team, alongside doctors learning on their own ([Cun95]). Recent research on learning in the workplace suggests that learning within workplace groups is very important but also highly dependent on the microclimate of that particular workplace (Eraut et al 1998). Both appropriate management training and contact with other groups can help in the process of improving the learning climate ([Bur98]).

The American researchers [FB98b] are also now advocating a co-ordinated approach to all three levels, suggesting that the role of CME providers should be to:

- Facilitate self-directed learning by providing for self-assessment, the acquisition of knowledge and skills, and the opportunity to reflect on clinical performance
- Offer high quality individual and group education that provides authoritative information, knowledge, and skills based on expertise and evidence
- Assist healthcare delivery systems to develop and practise organisational learning.

The developments reviewed above are recognised in the Department of Health's 1998 report, **A Review of Continuing Professional Development and General Practice**, whose principal recommendation is

> "to integrate and improve the education process through the **Practice Professional Development Plan (PPDP)**, developing the concept of the "whole practice" as a human resource for health care, resembling the health promotion plan in general practice and increasing involvement in the quality development of practices." (page 3)

The report anticipates the gradual fading out of the PGEA and argues that future CPD systems should "recognise and reward the process of need assessment, CPD planning and outcomes assessment" (page 14). It notes that further research will be needed

to ascertain how best to achieve this goal, and to link CPD with audit and R&D. Significantly it argues that:

> "the main message in delivering effective CPD is that the key to lifelong learning lies not in how to learn, but in how the learning process is managed." (page 13)

6.6 Summary

The relevant research into Continuing Medical Education and Lifelong Learning falls into three main categories: research into how doctors learn, evaluation of CME interventions and research into innovation strategies using single or multiple interventions to achieve changes in specifically targeted areas of practice.
Surveys of GPs, and also in a few cases consultants, have shown the importance for learning and changes in practice of a wide range of learning activities and sources of information. Moreover, they differ according to whether the changes involve treatment (including prescription), diagnosis and investigation, doctor-patient relationships, referral policy, health promotion or practice organisation. Models of physician learning distinguish between learning triggered by the problems raised by current individual patients and "learning projects" to acquire or improve proficiency in a targeted area of practice. The initiation of learning is dependent on significant background knowledge of what is out there to be learned to which CME conversations with other physicians, and reading contribute in ways which would not be revealed, for example, by evaluations of CME events.
Evaluations of CME courses have demonstrated the importance of including activities such as the observation and discussion of visual material and/or supervised practical work. Though it has confirmed that short courses of 1 day or less are rarely effective, no controlled studies have been reported which used length of course as a variable. This deficiency needs to be remedied because much time could be wasted trying to improve courses which are too short; and unrealistic expectations of the learning time required for certain goals are easily developed by busy learners and under-resourced providers - a form of collusion from which nobody benefits. Another important conclusion is that educational interventions on their own often fail to achieve changes in practice.
Research on innovation strategies points to the danger of focusing only on the development of competence. Competence has to be translated into performance and at this stage many dispositional and organisational factors come into play. Research on the implementation of guidelines, for example, indicates not only that the quality and utility of the guidelines themselves is important but also that both educational interventions (leading to understanding of their purpose and rationale) and administrative interventions (ranging from organisational changes to simple reminders) need to accompany the guidelines.

continued over page

Summary — continued

The discussion of recent developments in CPD reaches two conclusions. First, needs analysis is important for quality assurance purposes at three levels - the individual, the working group and the healthcare organisation (the last two are multi-professional). However, it should not be assumed that needs identified by audit, for example, will necessarily require an educational response. Second, following the advice of [FB98b], CME providers should adopt a coordinated approach to all three levels by facilitating self-directed learning, providing high quality individual and group education, and assisting healthcare organisations to develop and practise organisational learning.

Chapter 7

The Role of Information Technology

We can divide Information Technology into two main types. First, "Tools", IT to support doctors and dentists in the everyday conduct of their jobs; for example an MR scanner, a patient record system, or a decision support system. Second, "Training Systems", IT designed explicitly for training; for example a CD-ROM containing dermatology images and related text or a computer-based teaching system for a surgical technique.

Tools play a role in training not only because doctors need to be trained in their use but also because some tools can be exploited indirectly for training purposes, irrespective as to whether they incorporate a special training mode. For example, a decision support system can be used to help trainees reflect on their own decisions by providing an alternative analysis, not necessarily better or worse, to that of the trainee (some support for this notion is provided by [TKW92] who show that providing computer-generated corrective feedback of students' risk predictions of cardiovascular death based on the presence or absence of five factors, improved both the students' base rate calibration and discrimination). Moreover a database system can be used to store information from neurology department 'morning reports' that would otherwise be hard to find later, and then reuse this information for training purposes ([RKS95]).

As far as tools go the main educational questions concern

- The degree to which the tool assists or hinders practitioners:
  - the degree of fit of the tool to the work-setting in which it supposed to be used,
  - the way that the tool may redefine the nature of the task being done, and
  - the practical issues of installing, maintaining and providing appropriate institutional support (including technical) for the tool in order to provide an effective service.
- The degree to which new users of a particular tool are provided with the time and the assistance to learn how to make effective use of it. How easy to grasp is the interface of new tool, given the other tools that the doctor has met or needs to interact with on a regular basis?

For example, [KKP...96] describe a cognitive approach to evaluating a computerized patient record system and found that "two major classes of problems can be defined: problems associated with the user interface, and conceptual problems that arise when physicians try to map and represent findings with terms used by the system." (page 414). Similar issues of the effects of mismatch between computer-based representations (of clinical guidelines) and human expert representations are reported by [PAAS98].

As far as Training Systems go the main questions concern

- The extent to which the knowledge and skills of trainees actually improve as a result of the training system. Two important subsidiary issues here are
  - The degree to which the training tool provides a training experience that can be exploited by the learner in his or her work.
  - The degree to which the training system provides effective individual feedback so that the learners can track their own changing knowledge or performance.
- The degree to which the system provides cost (or time) savings for the same learning outcomes as traditional methods, for example, by reducing the time to reach some criterion of mastery.

IT in Medicine, as in other fields, throws up two distinct kinds of literature. There are papers which point to the latest technological advances — Virtual Reality, Multimedia, The Internet, Tele-medicine, for example — and anticipate many future benefits. Then there are more hard-nosed papers that evaluate training opportunities and outcomes using current IT technology in realistic settings.

In some ways this leads to a confusing picture with the reality often at odds with the promise. The situation is further confused by the varying outcomes of evaluations of computer-based training methods. However its clear that IT and medicine are inextricably intertwined and that the use of IT in training will increase.

In the first "anticipatory" category [MB97] provide a short, upbeat but useful introduction to some of the newer Information Technologies and [HV97] review some 40 Virtual Reality systems under development on anatomy and clinical skills teaching. Also in this category are the predictions for the "Class of 2003" by ([FE98]). They foresee the increasing delivery of "processed data", e.g. about treatment options, at the point of care delivery, and the increasing dependence of practitioners on validated and trusted databases rather than the primary (journal) literature. So they see a coming together of evidence-based medicine and clinical information technology to provide the practitioner with whatever information is needed at the time and place where it is needed. Of course this raises the issues of how the practitioner is to be trained to access and to *judge the trustworthiness* of such information and how systems are to be designed to make information available in an effective manner that meets doctors' needs ([EFC97]). In her introduction to a special issue of the *Journal of Artificial Intelligence in Medicine* [Pat98] argues that recent developments in collaborative technologies will cause people to view "cognition as a distributed process. In this perspective, intelligence, can be seen as distributed in designed artifacts such as computer-user interfaces; in representations, such as diagrams; and through communication in social contexts." (page 94). An example of this view is cited below ([SPC...98]).

In the same "anticipatory" category [OLS...95] suggest that a Virtual Reality (VR) based simulator linked to a fuzzy logic based evaluation system has great potential for the problem of training in laparascopic surgery. Likewise [DJ95] anticipate the benefits of VR in endoscopic surgery. Their paper provides a useful introduction to VR and its potential use in this kind of application.

By reference to artificial intelligence techniques in training areas other than medicine, [LK92] describe how collaboration and negotiation among differing viewpoints could be incorporated into medical training systems.

In the second "realistic" category, the survey conducted by [Kin96] paints a more depressing picture of the speed of the introduction of IT into the NHS out-pacing the possibilities of post-graduate medical education to provide sufficient training, particularly for practitioners. By contrast [CGdJ...94] provides evidence that high-fidelity computer-controlled simulators can provide effective training especially in those areas where there is a need to upgrade "competence in handling those uncommon but potentially fatal problems that require rapid and correct responses, without exposing the patient to risk". The domain they evaluated concerned problems during anaesthesia, but the argument can be made more generally. The positive evaluation of the learning took place *four months* after the anaesthetists had practised dealing with malignant hyperthermia on the simulator.

While more concerned with research than with training, [SPC...98] provide a detailed account of the way that five leading North American medical institutions are working towards a shared software infrastructure (InterMed) that will allow effective collaboration in many different areas, such as the development of guidelines, medical records, intergrated training, decision support systems and so on. The paper analyses the way that different technologies (e.g. email, conference calls) contribute to various kinds of collaborative decision-making and to some extent shows the ways things are likely to develop in the UK. An indication of the speed of change of technology (i.e. pre-Internet) is provided by [KKD...89]'s description of a computer-based information system for Post-graduate Medical Education in the west of Scotland, enabling online library services, computer assisted learning, word processing and statistics.

Although focused on medical students, [Fri96] lists ten reasons why the World Wide Web may not prove as useful in medical training as some would hope. These include the poor integration of CAI material into the curriculum, the lack of standards for judging the quality of CAI programs and the poor design of some, the failure to update the content of programs, insufficient access to computers, the mismatch between skills or knowledge taught and those needed and/or assessed, poor response times on the Internet. Addressing at least some of these issues, [MW96] provide a helpful account of how to make CAI work in practice based on experience at the University of Michigan Medical Center. Their paper also details a number of catalogues of medical training software.

Finally [LL98] provide a helpful, detailed and critical review of the application of artificial intelligence in medical education. They discuss the disparity between the promise and the outcomes of computer-based medical education in general, and artificial intelligence-based medical training systems in particular and describe a number of representative training systems.

7.1 Training Systems

The literature on training systems in medicine is large and scattered through a wide variety of sources including journals and books on computers and education in general, via journals on medical education through to journals for particular medical specialities. The following brief section simply provides a general indication of what is on offer.

7.1.1 Psychomotor Skills

Various systems have been developed both for assessing (see for example, [JHCS97, DFA...98]) and for developing psychomotor skills. These systems are designed largely for pre-registration rather than post-registration, but the lessons from these pieces of work generally apply more widely.

For example, [HRR96] evaluated a system for clinical clerks that provided a preoperative tutorial supplemented by a computer-based package for examining and interacting with anatomical images appropriate to the operation about to be performed. Their evaluation showed that the tutorial plus the computer-based system was beneficial in terms of the clerks' subsequent operating performance, but they did not attempt to assess the relative benefits of the tutorial, delivered by an expert, and the computer-based anatomy package. [MPF98] describe a computer-based system to teach minor oral surgery. The system offers multiple choice questions some of whose answers involve locating the correct point in an associated image or moving a cursor across an image, to indicate, for example, where to draw the needle when suturing. The evaluation of the system was in terms of what the users thought was good and bad rather than via an analysis of pre/post comparisons of increases in skill. [RRYH98] compared the efficacy of a computer-based system compared to a seminar in teaching 82 medical students how to tie a two-handed square knot. Although both groups learned to tie the knot, and could do so in similar times, the computer-based group produced poorer quality knots. A crucial difference between the two groups was that those taking part in the seminar received feedback from experts during the seminar, whereas the computer-based group received no feedback. This highlights the consistently hard problem for computer-based training of having the system monitor, evaluate and react sensibly to the learner's attempts to master the skill or solve a problem.

One partial way out of the feedback problem is to integrate the computer-based teaching system with "real" psychomotor skills laboratory. That way the students can compare how they perform suturing using, for example, reusable prosthetic skin, with multimedia video clips showing how it should be done ([OMKR98]).

An important issue, especially in surgery concerns the fidelity of the system and the degree to which skills learned with the computer-based system transfer to the real situation. Despite [CGdJ...94]'s positive evaluation in anaesthesia, cited earlier, [CRM...96] Chapman et al. (1996) found that for open thoracotomy assessment, performance of the task on a pig was a better discriminator of skill level than performance of the task on a computer simulation. They also found that practice with the computer simulation did not improve later assessment performance on a pig but did improve later assessment performance on the computer simulation.

7.1.2 Perceptual Skills

A computer-based system that makes no attempt at evaluative interactivity is a CD-ROM addition to a dermatology course ([HC98]). The CD-ROM supplemented an existing course using live-patient sessions and various sources of visual material. The CD-ROM added a 100 pages of text and diagrams. In a comparison with the course prior to the introduction of the CD-ROM, the medical students highly rated the CD-ROM but did no better in the examinations as a result of its introduction.

While there are many computer-based training aids for radiology, most are essentially electronic books or collections of images together with some kind of indexing mechanism, normally based primarily on disease. There have been relatively few systems that attempt to either model the domain or the evolution of knowledge and skill of the student in a detailed way, i.e. provide the evaluative interactivity lacking in more straightforward computer-based training packages. Of these, [AL98] describe an analysis of the problem solving operators used in mammography as applied by radiologists of various levels of skill. They also analyse the nature of teaching as it occurs in radiology case conferences and particularly the way that experts articulate their diagnostic reasoning. Both these analyses are used as part of the design process for RadTutor ([ALD...97]). A similar careful analysis in the domain of chest X-rays has been carried out by [Rog95a] as part of the design process of VIA-RAD tutor. [MMT...93] have taken a case-based approach in a tutor for CT and MR brain images. Their system offers a case-retrieval and decision-support mechanism based on descriptors. Their system also employs an atlas and contains tutorial material and images of normal brains as well as those displaying lesions. It can act as a decision support system by offering a range of possible diagnoses and access to the images of related cases, given the textual information that has been entered. [SdBT...95, SJT...97] have developed an image description training system that aims to help radiology trainees learn how to describe MR brain images in a systematic way by means of a structured image description language (IDL). This language allows clinically meaningful features of MR brain images to be recorded, such as the location, shape, margin and interior structure of lesions. The system is deliberately aimed to support and train the radiologist's inferences from what can be observed in the images.

An innovative problem-based approach is adopted by [KDH...98]. They get their students to develop their own multimedia packages (in areas of obstetrics and gynaecology). This provides the students with useful IT skills, including internet skills, as well as the chance to reflect on the topic chosen by building the package for other students to use. The authors note the problem of preventing over-enthusiastic students from spending too much time as well as the "significant time and effort to establish the programme".

7.1.3 Communication Skills

[HRJW97] describe a computer-based training system, INTERACT-CANCER, to teach communication skills, particularly those needed in dealing with cancer patients. Their paper offers a good pointer to the literature on communication skills and to related computer-based training systems. Their system consist of four modules. The first is a general introduction to the topic of communication. The second explains how to break bad news and the third is about providing information about treatment and future ex-

pectations. The fourth is about the emotional reactions of the patient. Each module offers video-clip examples of both good and bad communication practice and asks questions of the user to help them reflect on what they are seeing and hearing. What the system cannot do is to observe and comment on the user's own communicative competence (see the discussion of the knot tying tutor above: [RRYH98]). Their evaluation of the system concentrated on the perceived value of the system and how it was used but did not attempt to measure changes in communicative competence as a result of using the system.

7.2 Decision Support Systems

The use of Health Decision Support Systems (see e.g., [TS98] for a comprehensive overview) raises various issues in relation to competence and judgement. Under their earlier title of expert systems such systems offered much promise but did not have a dramatic effect on medical practice. Now there is a more realistic sense of their strengths and weaknesses, and some are in routine use. A comprehensive meta-analysis of the effects of computer-based clinical decision support systems is provided by [JLHM94]. Their paper shows how difficult it is to evaluate the effects of such systems in an unbiassed manner. Having started out examining over 700 papers they eventually analysed the results on patient outcomes and clinician performance of systems described in 28 studies. Within these studies they found only three with positive patient outcomes, and 8 without; however 15 had effects on clinician performance and 9 did not. They call for more research, particular for more research of the same standard as the blind, randomized controlled trials as used to support other health claims (though they acknowledge the methodological difficulties here).

A useful guide to decision support systems currently in routine use can be found at `http://www.coiera.com/aimd.htm`. This covers Acute Care Systems, Decision Support Systems, Educational Systems, Laboratory Systems, Quality Assurance and Administration Systems and Medical Imaging.

Their index for Decision Support Systems is presented below in Table 7.1:

Name	Status	Type	Entry Date
Dxplain	routine use	clinical decision support	Nov 7 1995
Epileptologists' Assistant	decomissioned	nurse progress note assistant	Sept 23 1997
Jeremiah	routine use	orthodontic treatment planner	Nov 19 1997
HELP	routine use	knowledge-based HIS	Jan 2 1995
Iliad	routine use	clinical decision support	Oct 23 1995
MDDB	routine use	diagnosis of dysmorphic syndromes	Mar 29 1996
Orthoplanner	routine use	orthodontic treatment planner	Nov 19 1997
RaPiD	routine use	designs removable partial dentures	Feb 9 1996

Table 7.1: Decision support systems in use.

The issue is no longer one of "can such systems be built and installed in medical care settings?", so much as "how useful are such systems in practice?" and "how does their use affect human decision making processes?"

A representative paper addressing the issue of the use of such systems is provided by [EFW...96]. Elstein et al (1996). They examined the effects of how using Iliad (see above) on four clusters of nine difficult cases affected the decisions of 16 doctors of various levels of experience. Each doctor dealt with nine cases. Their main finding was that Iliad produced a list of diagnostic possibilities containing the correct diagnosis in 38% of the cases. This success rate was worse than the most experienced doctors (43%) but better than for residents (33%) and for fourth year medical students (15%).

Each physician worked on each case both with and without Iliad. Working with Iliad produced improvements in the diagnostic accuracy of 15% of the cases dealt with by experienced physicians and by the medical students but no improvements in cases dealt with by residents. On the negative side, in about 12% of the cases dealt with overall there was a decline in accuracy, with the largest decline among the residents (statistical significance not reported).

Methodologies for assessing how decision making is affected by the use of systems such as Iliad are explored by [KP98]. They report a variety of effects such as problems in navigating through such systems and, more importantly, shifts in the manner of conducting patient interviews from their natural strategy towards conforming exactly with the menu of questions on the screen of the system — a "screen-driven" strategy. They also found that the differences in reasoning strategies between novices and experts (see Section 3.1) meant that the design of such systems needed to be adjusted to take the nature of the user's expertise into account in order to reduce the chances of mismatch between the line of reasoning of the system and that of the doctor. A similar IT-based mismatch issue is described by [CP98] who outline the problems that patients have when using a "telephone-based telecommunications system" to describe their symptoms to a remotely located doctor. Analysis of the interactions showed that the organization of the information required by the system suited the doctor's way of thinking but did not suit the patent, with consequent errors of communication.

[TS98] identify the the major changes in direction that decision support systems need to take as follows:

- "a shift in focus from fragmented data to linked databases
- increased emphasis on the development of more powerful models and improved knowledge engineering methods
- increased emphasis on research to determine and explain decision outcomes" (page 376)

The latter point links back to [KKP...96] above with the need to ensure that physicians can understand how the decision support system has framed the problem so that they can judge the quality of the advice that is being proffered.

7.3 Summary

A great many systems have been developed for various aspects of training, but most come up against the the consistently hard problem of having the system monitor, evaluate and react sensibly to the learner's attempts to master a skill or solve a problem. Decision support systems have had a mixed reception over the years, but seem to be growing in acceptance as just another tool in the doctor's armoury. Researchers foresee an increasing delivery of "processed data", e.g. about treatment options, at the point of care delivery, and the increasing dependence of practitioners on validated and trusted databases rather than the primary (journal) literature. They anticipate a coming together of evidence-based medicine and clinical information technology to provide the practitioner with whatever information is needed at the time and place where it is needed.

Chapter 8

Assessment and Revalidation

Assessment is not the prime focus of this report, but it plays an important part in the development of competence and judgement. Three types of purpose can be usefully distinguished:

- Certification or qualification of individual doctors
- Provision of information to guide learners and/or those who supervise and support them
- Quality assurance and improvement of practice.

Certification and Qualification, particularly when examinations are involved, are often described as "high stakes" decisions. For examinations, the risk to candidates will depend on pass rates, the possibility of reassessment and the probabilities of first time and second time success. The cost of failure will depend on whether their career progression is significantly affected, as well as the need to engage in further training or study. However, even for candidates least at risk, the effect of any high stakes assessment is a considerable increase in attention to what is being assessed , with a concomitant decrease in attention to what is not being assessed. Hence the quality of an examination has to be judged not only by whether the right people pass (false positives being of particular concern whenever certification is involved) but also by its effect on candidates' direction of their learning effort. This washback effect of examinations may be positively valued; [WS92], for example, report how the introduction of a Critical Reading Question paper into the Membership Examination of the Royal College of General Practitioners increased time spent by candidates on critically discussing papers and reading two key journals. Where negative views are expressed, it is usually because the examination is thought to overemphasise knowledge at the expense of competence (see Chapter 4), to neglect a holistic approach by focusing mainly on components of competence rather than their integration, or to emphasise competence at the expense of performance (candidates exhibit their best behaviour under ideal conditions).

8.1 Assessment Methods

Excellent overviews of assessment methods used in examinations are provided by [vdVN94], [SNF95], and [FB98a]; and advice on the development of specific tests by [New94, DFJ...94]. Although all the methods reviewed are used at postgraduate level, much of the published research refers only to undergraduates. However, some development are clearly relevant at both levels. For example, [PB95] have taken advantage of recent research on clinical decision-making, supporting the view that problem-solving skills are case or problem specific rather than generic to replace written Patient Management Problems whose validity was increasingly in doubt ([NBC...85, BP87]) by sets of problems focused only on those **key features** that are crucial to their successful resolution. Not only does this cover more cases in less time but it achieves high validity by adopting a more holistic construct of competence.

Several different groups have reported developing Objective Structured Clinical Examinations (OSCEs) for use at postgraduate level: [FMM94] for GP Consultations; [SDJ...93], [SDS...95] for surgical interns; [HRHM98] for psychiatry. [HCBN96] paper on using OSCEs specifically targeted at difficult communication skills suggests that these also have a strong case-specific element, with consequent implications for going beyond generic training in communications. Another development in Canada involved lengthening the time spent at each station on a multi-case examination to 30 minutes in order to make a more comprehensive examination of the more sophisticated knowledge of senior surgical residents ([MCR...97]). The resulting Patient Assessment and Management Examination (PAME) had better psychometric properties than other measures, when using 6 cases. Patient satisfaction ratings were included with global ratings by the examiner for each of the four phases: - initial patient assessment (8 mins), ordering and interpretation of investigations (4), a second interaction with the patient to discuss diagnosis and management (10) and a structured oral examination (6). A parallel development in Ireland led to a objective structured long Examination Record (OSLER) in which the candidates spend 20-30 minutes with the examiner alone, having already been observed examining and taking a history from the patient ([Gle97]). Including a sufficient number of cases is still important for reliability; because case content is the most significant variable ([vdV96]) and results should not be left to "the luck of the draw".

Both OSCEs and PAMEs/OSLERs require the use of simulated standardised patients ([vdVS90]), though [Bar93]'s review points to other, less structured use of standardised patients to assess a doctor's general approach to clinical examination rather than specific skills. [SSR...91] estimated that 2 half-days interacting with 19 Standardised patients was sufficient for evaluating the data gathering and interviewing skills of residents.

[PTODM94] found that GP trainees performed better with actor-patients than in recorded consultations with real patients in daily practice, but this evidence nevertheless provided good predictions of weak performance from trainees. They interpreted this as a distinction between competence and performance rather than lack of validity when using standardised patients: [FSK95] used actor-patients for testing communication skills at the end of a Diploma course in Palliative Medicine; and showed that, although their ratings of the doctors were a little higher than those by the "official" examiners, they correlated well. Indeed standardised patients have been trained to give formative feedback to medical students and can play an important role in this aspect of their training

([Bar93, SRPH90]).

[RRM...97] and [MRR...97], extended the use of OSCEs to test technical skills with their objective structured Assessments of Technical Skill (OSATS) for measuring surgical residents' technical skill using bench model simulations outside the operating room. [JSM...96] have developed a test for technical clinical skills of general practitioners in the CME context. A more sophisticated but not highly expensive simulator (£3000) has been used by [BJ97] to assess anaesthetists responses to a range of simulated emergencies; and further work showed that under simulated critical conditions chart recording errors markedly increased ([BSJ98]), thus raising questions about the accuracy of records of such incidents in real situations.

Although associated with formal examinations, these assessment techniques can also be used for formative purposes, to provide feedback to the trainee and/or trainer. Thus [PTODM94] used assessment with simulated patients to identify trainee GPs needing more support. [Gle97] reported how immediate feedback from an OSLER led to a significant improvement in performance. [SDS...96] found that feedback could be given to participants during an OSCE without perturbing test reliability. [GTW...94] describe the increasing popularity of a voluntary in-training examination for residents in internal medicine. 47% took this opportunity to compare their performance on a comprehensive written examination with national norms, of whom 45% were second year residents.

8.2 Assessment of Performance

The term **performance assessment** is often applied to any assessment involving real or simulated patients; but, mindful of the important distinction between competence and performance, we shall refer to all the methods discussed above as either competence assessment or assessment of knowledge. The term performance assessment can then be confined to assessment of real on-the-job performance under working conditions. Such assessment is the only form of assessment under a "pure" apprenticeship system and has always played a significant role in the UK. [Dau95]'s review of the increasing use of performance assessment in North America supports the trend towards improved validity but also analyses the reliability problems of new approaches. The least satisfactory of all formal assessments in the UK must surely be that associated with admission to the register. Though subject to ratification, the "satisfactory completion" of a doctor's pre-registration year depends on global judgements by their supervisors, made with varying degrees of rigour and varying amounts of unaggregated evidence.

Registration as a general practitioner now involves a formal summative assessment based partly on videorecordings of patient consultations. Since these are selected by the candidates as representing their best practice, they are sometimes described as assessments of competence rather than performance. A pilot study in Scotland ([CHM95]) concluded that assessors reached firm judgements by the time four consultations had been viewed, and estimated a 95% probability of identifying a non-competent trainee. It also includes a trainer's report informed by a checklist which is an overall rating of performance and takes into account dispositions and attitudes as well as clinical competence, a written examination based on Multiple Choice Question (MCQs) and an audit marked against set criteria.

The newly introduced Certificate of Completion of Specialist Training (CCST) represents a major change in policy for hospital doctors. Regional training programmes with built-in rotations are the responsibility of the postgraduate dean, together with the annual progress review and final assessment of specialist registrars (SPRs). The annual review is conducted by members of the regional specialist training committee not directly responsible for the trainee on the basis of inspecting each trainee's logbook and cross-examining them on it contents. For final assessments 1-2 specialists from outside the region are added. The logbooks are designed by the appropriate Royal College and specialism, and typically include a list of competencies, each of which can be assessed at several levels. Two examples are shown below for the Royal College of Obstetricians and Gynaecologists and the Royal College of Physicians. Entries are made jointly by the registrar and supervisor.

The RCOG levels of competence range from observation (1) to independent practice (5). The following list specifies what is meant by each level.

Level 1	Observes	Observes the clinical activity performed by a colleague
Level 2	Assists	Assists a colleague perform the clinical activity
Level 3	Direct Supervision	Performs the entire activity under direct supervision of a senior colleague
Level 4	Indirect Supervision	Performs the entire activity with indirect supervision of a senior colleague
Level 5	Independent	Performs the entire activity without need for supervision

Level 1 (observes) and Level 2 (assists) include the presentation of basic and clinical knowledge, exhibition of clinical reasoning and identification of relevant principles associated with the target activity.

The RCP levels are

Level 0	Insufficient theoretical knowledge
Level 1	Theoretical knowledge but not competent
Level 2	Some competence
Level 3	Fully experienced and competent

Some colleges and specialisms also include written examinations in their requirements for CCST, though most regard their Membership Examinations taken before entering higher specialist training as providing a sufficient foundation of basic knowledge. Further knowledge, especially of recent research, is assumed to be included within the assessment of clinical performance.

Until recently, the traditional practice of using global ratings by supervisors as the main indicators of satisfactory progress has been a major weakness in the UK system of specialist training; because this approach has been repeatedly shown to suffer from halo affects, evaluator leniency and restricted use of the grading range ([NBC...85]). As a result weaknesses in performance which ought to be picked up and dealt with are allowed to continue with adverse effects ([LT97]). The use of other methods, a wider range of raters, improving the rating instrument and training the assessors ([WSW95]) are all advocated as remedies, preferably in combination. Thus the quality of assessments in the new SPR logbooks will require careful monitoring and the will to take further action if necessary. The most obvious extension would be periodic structured ratings of observed

performance using purpose-developed rating scales (e.g., [WRCT94]).

Given the importance of progressing doctors more rapidly through what is now a shorter but more structured training programme, there is a need for some form of progress chasing which also improves quality of performance. The credentialing of internal medicine residents for undertaking procedures ([GBI91]) was discussed in Section 5.1. Another approach developed by [BN97] in Obstetrics used "encounter cards" with scales for knowledge, professional skills, manual skills and overall performance. At periodic intervals, residents were asked to have 6 to 8 encounters scored during a particular week (reducing the selection bias) by faculty who observed them with the patient. Immediate feedback was given, and the full sequence of cards reviewed a little later. This process is less unwieldy and probably more valid than an examination, and its purpose is purely formative. An alternative strategy, not requiring observation but more focused in its target, is to review the case notes of the last 10-15 patients with a particular condition to be managed by the trainee ([Pie93a]) (Pietroni 1993).

A generic quantitative approach developed in Scotland ([MGAC96, PGAC96]) uses a workload system, which weights cases for complexity and degree of supervision to estimate the effective operative experience of surgical trainees. [ASM...89] suggest that for some pre-surgical candidates performance outcomes can be used as indicators of clinical acumen. Thus residents were asked to predict the likelihood of 101 patients having appendicitis, while evidence of outcomes was collected from pathological inspection after appendectomy or records of the alleviation of abdominal-pain without appendectomy. This enables the construction of a Diagnostic Ability Score (DAS). A similar approach can be used for regularly performed procedures, such as colonoscopy (for surgery) or central venous cannulation (for anaesthetists); and made more flexible by using cusum analysis, a cumulative average failure rate. This gives both (1) an objective account of trainees progress for planning further training or taking remedial action and (2) a cut off point for credentialing ([WPS92, vRMP...95, Kes95, Kes96]).

While examinations can and should provide feedback to candidates and trainees to enable them to adjust learning goals and the provision of learning opportunities, the link is usually close for performance assessment. Through most of the papers reviewed above, there is a major concern for assessments which provide feedback appropriate for remedial action by the various stakeholders; and the corollary is sufficient commitment to remedial action to justify not just conducting assessment but implementing them with a disposition and skill that preserves their effectiveness and ensures their contribution to the quality of training. These considerations are even more important when we attend to the problems of developing and maintaining the competence of doctors who are already qualified. This may be in the context of quality assurance, needs assessment and/or evaluation of CME programmes and policies, or recertification/revalidation.

[Pol96]'s review of measuring the performance of established surgeons covers communication, mortality and morbitity meetings, CME , peer review and audit. Except for communication these are context categories, the evidence used may come from statistical monitoring, a thematic review of a particular category of cases, or a review of individual cases. Local statistical data is available from medical records, but often additional data are collected for thematic reviews. Their utility is improved by the increasing availability of regional or national norms or relevant research findings linking data with patient outcomes. Such evidence is more clearly attributable to individuals in surgery than in

most other specialisms.

[LH94]'s analysis of chart review noted its value in collaborative efforts to improve patient outcomes in a particular community; but also reported its limitations as the sole method for assessing a physicians's competence. The interpretation of chart data can be variable so a large number of cases (or judges) are needed to establish reliability. Whereas audit may be a useful trigger for quality improvement, it is too blunt an instrument to provide reliable assessment of individuals unless they show extreme deficiencies ([MPM...90, NMHH93]).

8.3 Recertification or Revalidation

The Canadian repsonse to this problem has been to develop a three tier system for the monitoring and enhancement of physician performance which is expected to become mandatory. This involves:

1. Initial screening of all physicians (using fee for service billing patterns, peer assessment questionnaires and patient satisfaction questionnaires) on which feedback is provided ([PBD...95, JSB...95]);

2. Assessment of about 10% of practitioners identified by the first screen as being at risk, (using hospital or office audit and structured interview) followed by appropriate CME programmes for groups and individuals;

3. Detailed assessment of high risk physicians- (1-2%) by a mixture of methods, followed by a specific remedial programme and possible removal from practice.

Assessment techniques appropriate for the third stage were reviewed by [NDL...93] in connection with a similar programme operating in Ontario for Primary Care Physicians. Their 7 hour package, incorporating chart-stimulated recall from their records, an oral examination, of 3 cases, 3 standardised patients, a 5 station OSCE and a Multiple Choice Examination, proved to be highly reliable and fit for its purpose. The Ontario programme has been screening a random sample of 450 physicians a year, of whom 0.5% eventually required removal from practice.

In the US doctors have to get recertified every seven to ten years if they want to retain their membership of a specialist Board (the US equivalent of a Royal College). This does not affect their license to practice medicine in general, but their right to claim specialist status and fees. The process entails passing an examination which is increasingly becoming computer-based. Although based mainly on propositional knowledge, its importance should not be underestimated.

[MPM...90]'s survey revealed disturbing evidence about some doctors failure to keep up-to-date; and [TARB97] showed that GPs self-assessment of their level of knowledge was highly unreliable. The scope of these examinations is also likely to increase as the technology enables it to cover various forms of simulated practice. However, on-site review of practice has recently been discontinued because of the expense; but their cost would be much less if they adopted the Canadians; three tier system whereby only 10% of those doctors screened receive later on-site visits.

British policy is currently in a state of rapid change. [SD98] report the progress of the first new policy to be implemented, the performance assessment of practitioners screened by the General Medical Council (GMC) as the result of a complaint. If the issue is one of poor performance then a peer review procedure not unlike the Canadian second stage is triggered. This entails a visit by two medical and one lay assessor from the GMC who engage in the following activities:

- Assessment of medical record keeping
- Discussion of the management of the doctor's own cases and clinical work
- Observation of aspects of actual practice
- Audit of clinical outcomes
- Interviews with third parties
- Structured interview with the doctor
- Site tour to determine the circumstances of practice.

If the visit cannot rule out a serious deficiency of performance, then it is rapidly followed by a second stage involving formal tests of competence: - a written or oral test of practice related knowledge and clinical thinking, a test of consultation skills (clinical thinking and communication skills), and a structured clinical examination of practical clinical skills.

Mindful of increasing public concern about the competence of a significant minority of doctors and the small number of evidence-based complaints reaching them, the GMC recently announced proposals for the **revalidation** of specialists and general practitioners, which they intend in due course to extent to all doctors (GMC, 1999). The term "revalidation" appears to have been chosen to signify a more normal, less threatening event than "recertification". The central policy problem, given the inevitable cost of such a process, is to achieve three separate goals:

1. To detect, and where possible remedy the deficiencies found in, incompetent doctors;
2. To advise doctors about weak elements in their current practice which might lead to incompetence if not remedied, thus forestalling a more serious verdict at a later date; and
3. To contribute to the professional development of doctors whose competence is not in doubt.

The first of these goals is a reasonable proposition given the GMC's recent experience of performance assessment and recent developments in Canada. The second is more problematic. Although [Irv97] emphasises the need for "sound local arrangements for recognising dysfunctional doctors early and for taking appropriate action" (page 1613), [Don94]'s review of 49 problem cases concluded that "existing procedures for hospital doctors within the NHS are inadequate" (page 1277).

However, this is now being given high priority at local level as the fourth main component of the new policy on **Clinical Governance** ([Nat99]). The Standing Committee in Postgraduate Medical and Dental Education (SCOPME) recommends a combination of clear standards and peer review, recognising that each is in need for further development ([Sta99]). As implied by our earlier definition of competence (Chapter 2), the standards must reflect the expectations of a particular doctor in a particular job. They cite the example of the self-assessment manual and standards of the Faculty of General Dental Practitioners and its use by individuals and voluntary peer review schemes. Given its brief, SCOPME's perspective is primarily educational and they have argued for some time that CPD, review and support should not be linked to regulatory or disciplinary arrangement. Thus they are concerned that: "revalidation could easily be seen as a threat and continuing registration could become the sole focus of clinicians' responses, utilising their energy and available resources at the expense of their professional development" (page 14).

SCOPME rightly argues for more research and development into the causes, assessment and educational remediation of poor performance; but nevertheless conclude with a tinge of euphemism that: "the main challenge for everyone will be to find ways of making the revalidation process both a valued part of professional life and, for the overwhelming majority of doctors, an important stimulus to further development" (page 14).

The factor ignored in this debate is the greater frequency of annual reviews associated with Personal Development Plans and Practice-Based Development Plans. If these incorporate an element of peer review (see [Bou99] for a description of the Dutch system), then periodic revalidation every 5 to 10 years should not weaken personal commitment to CPD. Indeed, if followed by a universal entitlement to a period of educational leave with a mutually agreed focus, achievement of the third goal of revalidation would be greatly enhanced.

8.4 Summary

Hitherto, most of the research effort has focussed on the assessment of competence linked to certification decisions rather than the assessment of performance on-the-job. This emphasis is gradually changing as public demand for robust quality assurance grows. Revalidation is about to be developed with very close attention to research on performance assessment. A range of assessment methods is reviewed, and the unreliability of the traditional practice of using global ratings by supervisors noted. Persistent conclusions from assessment research are the need to use several methods, to refine all instruments or protocols, to train assessors and to use several assessors.

The Canadian three-tier system for the monitoring and enhancement of physician performance is now well developed and familiar to those exploring revalidation in the UK. The need for good assessment practice linked with effective strategies for the improvement of practice (see Chapter 6) is critical, both for revalidation and for formative and summative assessments during postgraduate education.

Chapter 9

Implications for Policy

9.1 Introduction

This chapter examines the implications for policy of the literature reviewed in this report. Its arguments are more personal because making links between research and policy in educational and social settings often involves extrapolating the conclusions of research beyond the context in which they were obtained; and this entails a change in the status of one's comments from proven conclusions to potentially useful insights. Since there is a substantial amount of North American research in areas relatively unexplored by the small number of British researchers, it is important to investigate the implications for British policy and practice without assuming that the transfer of findings across the Atlantic can be taken for granted. A brief introduction to the US system of residency training after graduation from medical school is provided by [Sal95].

Although we have organised this discussion on policy implications chapter by chapter for ease of reference back to the relevant research, we have not stuck to this format rigidly but have introduced (with an appropriate cross-reference) material from other chapters when it improved coherence and avoided two separate discussions of essentially the same issue. A recurring conclusion throughout this chapter is the need for further research. Descriptive data about the processes of British postgraduate medical education is sparse and outcomes evidence, other than pass rates for the Royal Colleges' examinations, is very rare indeed. More attention has been given to research into undergraduate education which is generally more accessible to researchers. The overwhelming busyness of the postgraduate experience is not only a constraint on the pursuit of its educational goals but also a deterrent to research.

9.2 Competence and Judgement

Postgraduate trainees are employed as working health professionals, so their current competence is always an issue, as well as the competence they will be expected to demonstrate when their specialist training has been completed. Defining competence in terms of the expectations of the holder of a particular post helps to avoid confusion; and it should be possible to make periodic updates of a doctor's progress beyond the minimum competence for the post to enable responsibilities to be extended to match their devel-

oping competence without putting patients at risk. Many log-books seek to achieve this but their usage needs to be evaluated. The central problem of postgraduate education is how best to combine work within a doctor's current competence, itself a source of learning, with the provision and use of learning opportunities to extend that competence. Training programmes are designed with clinical experience as the prime consideration, but many factors can affect the use of that experience for learning purposes. These include time to think, timely help, and an ability to discern and pursue those learning goals that should have highest priority at that stage of the doctor's development. Self-directed learning requires a strong focus as well as learning opportunities. Hence there is a need for an agreed framework for the development of competence during each training programme, which puts learning outcomes on the priority list **alongside service duties**. While many experienced clinical teachers may have an implicit framework for progression, this is insufficient to provide the necessary coherence and continuity. The distributed nature of learning opportunities and support in most healthcare settings requires the use of an explicit framework for the development of competence. Moreover the process needs to be efficient as well as effective if more is to be learned in the limited time available for training.

To develop such progression frameworks in a manner that recognises the enormous complexity of the task will need further research into the specification and communication of the competence required at various stages of training, at the completion of training and for progress beyond that point. It then becomes possible for training programmes to indicate how they provide learning opportunities appropriate for the expected outcomes of a particular training period. The recent report of the US Federated Council for Internal Medicine Task Force, Graduate Education in Internal Medicine: a Resource Guide to Curriculum Development ([SEKR97]) is a useful source of ideas. Not only does it set out a detailed framework of competencies which covers the whole spectrum of the physician's role, but it also discusses the advantages and limitations of different learning settings and provides matrices for mapping those settings where particular competencies are intended to be developed. Clarification, at least at a broad level, of what and how a doctor is expected to learn over a particular period is necessary for auditing and ultimately for researching the educational process. Until this happens the relationship between potentially available learning opportunities, received learning experiences and learning outcomes will remain obscure.

Areas in which competence is seriously underconceptualised include communications, teamwork and management in healthcare settings. The blanket term "communication skills" is used to cover a wide variety of processes and settings with little attempt to differentiate between them, to take into account the situated nature of communication or to develop frameworks for progression. Given the increasing priority being given to client-centred practice and research evidence demonstrating the impact of good communication on health outcomes, much more detailed and professional attention to this area is overdue. Communication between doctors and with other professionals is also critical, though not so often recognised as an area where a focus is required on what has to be learned and how. Teamwork and other organisational factors also impact on patient outcomes; but there is little evidence on how (or even whether) these capabilities are developed during postgraduate education. It is also one aspect of the concern expressed about the potential gap between the competence required to be given a CCST and that required of a consultant.

The capability to convert the competence of a range of professional workers into a team performance that meets expectations has to be developed by good supervision and management. This requires sensitivity to organisational factors affecting other professionals' performance as well as one's own; and the disposition to seek changes where these are needed for health care improvement. All these interpersonal foundations need to be laid down early in professional life, a responsibility which has to be shared and coordinated among all those carrying educational roles.

The inclusion of judgement in our brief has had a liberating effect, because there is considerable debate about whether it is , or even can be, included within the term competent. Some perceive judgement as an attribute of personal expertise that goes beyond that competence which any fully trained doctor could be reliably expected to demonstrate. It can also be seen as a dimension of lifelong learning linked mainly to the improvement of decision-making through learning from experience over a long period, rather than the learning of new practices or keeping up to date with research. Judgement is associated with complexity and uncertainty; and people find it easier to cite examples than to define it. Probing examples to elicit the nature of the underpinning knowledge is difficult; and the development of judgement by doctors has been little researched. We would expect in-depth discussion of difficult cases to contribute to such development, providing there was a learning intention.

9.3 Theories of Expertise

Theories of expertise developed in different contexts using different research techniques may emphasise different aspects but do not greatly differ in their conclusions. Key features include the importance of case-based experience, the rapid retrieval of information from memory attributable to its superior organisation, the development of standard patterns of reasoning and problem-solving, quick recognition of which approach to use and when, awareness of bias and fallibility; and the ability to track down, evaluate and use evidence from research and case-specific data. Understanding the nature of expertise is important for self-monitoring one's use of heuristics and possible bias, sharing knowledge with others and supporting other people's learning. It is also critical for understanding the respective roles of clinical experience and evidence-based guidelines. Those responsible for developing, disseminating, evaluating and modifying guidelines, decision aids, information systems and communications aids within teams and across teams need to match their procedures and modes of representation to the way doctors' minds work.

Research into decision-making under conditions of stress and uncertainty suggests that training in crisis management is needed, and that teamwork and other organisational factors are important. At the individual level there is a need to accept that this will always be a problem area and that non-cognitive factors are important, with confidence also being a critical aspect of performance. There is a need for regular self-evaluation to maintain critical control of one's practice.

Implications for teaching are the advantages of basing progression frameworks on case typicality and coordinating the use of exemplar cases and generalisable knowledge. The use of recorded material and standardised patients for developing competence in communication skills is strongly supported by research but the conclusion in Section 9.2

above was that there is still no evidence of any long term strategy for broadening and deepening this competence. The use of evidence-based medicine requires on-the-job as well as off-the-job teaching.

9.4 Implications for Training

The literature on expertise emphasises that it is the structure of experts' knowledge as much as the quantity of their knowledge which defines their capability. Clinical reasoning is optimised in the expert's particular area of expertise. Such reasoning tends to be 'schema' driven and differs from problem to problem ([NTF...85]) rather than following some consistent, classical model of the hypothetico-deductive approach.

These findings have implications for training. For example, [RN96] make the following suggestions (in relation to medical schools, though many of the issues apply post-registration):

Organization of memory: "...information in isolation is inert and unhelpful. ...it is inappropriate to try to teach clinical-reasoning skills independently of clinical content."(page 992)

Storage and retrieval from memory: training goals should attempt to "enhance meaning, to reduce dependence on context and to provide repeated practice in retrieving information." (page 995)

Problem solving: "The work on specialization of routines suggests that individuals must practice the same problem-solving routines regularly. The work on analogy suggests that, in addition, the novice must have intensive practice in identifying the situations in which a particular problem solving routine is likely to be useful [and *not* useful]" (page 997)

Concept formation, categorization, and pattern recognition: "...experience after graduation and licensure is of crucial importance to expertise, and it forces us as educators to more critically examine the relative contributions of analytical "textbook" knowledge and experiential knowledge to professional expertise." (page 999)

Decision making: "One implication from the literature on heuristics and biases is that medical students should be taught about biases, admonished to be aware of their pervasive nature, and be trained to avoid them. We are not aware of any evidence that such error-correcting training can be successful. In fact, the literature on the context-dependence of problem-solving strongly suggests that such general strategies toward caution are unlikely to be successful. Perhaps rather than focusing on "debiasing," instruction should be directed at helping students recognize those relatively infrequent situations where their heuristics are likely to fail.

A second implication is that medical students should be taught to make extensive use of decision-support systems, which are designed to be immune to decision biases." [RN96]

acknowledge that if the data collected and input to the system is subject to bias, then so will be the output. So they go on to say that the aim should be to help "individuals decide when the decision-support system could be beneficial rather than teaching individuals to rely on it extensively". (pages 999–1000)

An alternative to the decision-making aids approach is expressed by [Gra89] who argues for the employment of a course that helps doctors reflect on their thinking processes via a "large series of exercises which work with the participants' thinking processes *as they are*" (our emphasis).

One possibility is simply to ensure that those being trained are exposed to a sufficient number and variety of cases to allow them to build up the appropriate schemas. Another is to provide cases plus "instructional road maps" to traverse them ([FCSDS92]). Another approach, e.g. as advocated by [MJWH97], is to teach clinical problem solving using schemes. Of course, "using a scheme" is not the same as an expert "having a schema". Their approach (at Calgary) goes beyond simple "problem-based learning" by helping students develop specific schemes for each type of presentation (see also, [PWC...90]) in the area of expertise — as opposed to a more general methodology of generating and then refining multiple hypotheses. This view is disputed by [PSA96] who argue that some of "case-specificity phenomena might be viewed in part as an artifact of an educational system containing widespread inconsistencies in the instruction or assessment of disease class-specific differential diagnostic diagnosis concepts." (page S12).

An obvious question to ask is whether instruction in evidence-based medicine leads to improvements in practice, and then whether those improvements to practice themselves lead to better patient outcomes. [NS98] reviewed ten studies (from 1966 to 1995) that examined the effects of teaching critical appraisal skills. After excluding studies with methodological flaws, they analysed four studies of students and three of residents. They found that:

> "...although instruction in critical appraisal (evidence-based) skills can result in sizeable gains in knowledge among students, the effect of such instruction is much smaller among residents. Furthermore, the minimal evidence to date does not, as yet, provide any indication that the gains in knowledge result in a change in behaviour with respect to the critical use of the literature." ([NS98], page 180)

In pondering the differences in outcome between students and residents, [NS98] suggest that this may be to do with the different degrees of integration of the instruction into the educational program. For example, some of the students would have taken the course for credit, whereas the residents worked to a 'journal-club' format. They suggest that more positive effects may be found when such programs are more strongly integrated into all aspects of training (though they point out that they found no evidence to back this optimism). They also concluded that they could (as yet) find no convincing evidence that the "gains in knowledge demonstrated in undergraduate critical appraisal courses can be sustained into residency and practice and eventually translated into improved patient outcomes" (page 181).

Basic science and clinical science can be usefully seen as parallel fields of knowledge which illuminate each other but do not necessarily determine each other. Opportunities

for using scientific knowledge are often neglected, because people fail to recognise how much further learning is involved in transferring knowledge from an academic context to a clinical context. The selection and timing of science-based inputs to postgraduate education should be planned from a user perspective; and where appropriate the use of scientific knowledge be taught through case discussions in clinical settings.

9.5 Learning in Clinical Settings

There have been several evaluations of postgraduate basic training programmes in the UK. Though there have been a few improvements, the overall impression is still negative. Many features of the educational policy seem to be appropriate, but they are not being implemented in many hospitals. There is insufficient supervision and feedback. Educational goals are subordinated to service demands. While many house officers receive good clinical teaching, a minority do not and assurance of educational quality is weak. Learning goals are only specified at a very general level, so there is little clarity about priorities, especially at the PRHO stage.

Since service responsibilities contribute greatly to the development of competence, working and learning will often be indistinguishable activities. But they still signify different expectations of doctors in postgraduate education, and the tension between their respective priorities is constantly noted in both research studies and policy reviews. The educational dimension cannot easily be sustained by a *laissez faire* approach which allows problems in responding to today's patients to assume precedence over those of tomorrow's patients. This issue has to be tackled at local level where there is limited management of the educational process and clinical tutors have little time and no authority over clinical teaching. Deans do what they can; but quality assurance of postgraduate education lags well behind that for clinical practice; and the UK research base at this level is minuscule.

Another important issue which equally affects education and service goals is continuity of care ([Irb95]). Lack of opportunity to follow patients over time can easily become a hidden weakness in junior doctors' experience within the hospital setting, in outpatient clinics and across the boundary between primary and secondary care. The same principle can be applied to the junior doctors themselves, for whom short rotations limit the development of relationships and weaken the scope for the kind of supervision which incorporates wider aspects of the professional role and facilitates a learner-centred approach.

> Short clinical rotations require both teachers and learners to quickly determine how to work together for the care of patients. From a systems perspective, this loose connection between teachers and learners tends to inhibit close supervision, reduce targeted teaching, and limit thoughtful feedback. In addition, it increases the difficulty and complexity of the tasks of teaching and patient care. Mutual knowledge of the participants in the process (patient, learner, and preceptor) enhances the quality and efficiency of the interaction as well as the satisfaction of the participants. ([Irb95], page 906)

Continuing concern has been expressed about the survey evidence on basic surgical

training. Sometimes the problem is too little supervision of operations by house officers, sometimes the house officers get insufficient clinical experience. Operating under supervision is seen as the most critical feature of learning to be a surgeon and there is not enough of it. Some authors recommend the greater use of 'skills labs' and simulators. (See also Chapters 3 and 5).

The learning of procedures in medical posts has been criticised for being too haphazard: there is often little continuity of experience and guidance is often provided by doctors who are themselves not very experienced. More planning could enable more systematic teaching by more expert doctors; and a credentialling system like that used in many American hospitals would improve the quality assurance.

The appropriateness for GPs of so much general hospital training has been questioned. Though research on this issue would be difficult, we think more research evidence could and should be gathered. One year's training in general practice seems very short, especially since research in Holland resulted in them extending this period to two years. Not surprisingly research investigating particular areas of expertise has resulted in long lists of needs for GPs continuing medical education (see Chapter 6). In areas such as palliative care and psychiatry the argument seems particularly strong, in other areas one might look to other ways of distributing more specialist expertise within primary care organisations. GP training contrasts favourably with basic hospital training in its ability to provide tutorial support on a regular basis. Its quality varies no more than clinical teaching in other settings; and the commitment to quality improvement appears to be much greater.

Outside general practice, there is much more in-depth research in North America than in Britain. This has given particular attention in the last decade to learning in ambulatory care settings, a term which covers both family medicine and hospital clinics. More use is now being made of such settings in order to give doctors a broader experience of medicine, especially when significant aspects of care are being moved out of hospitals. The key issues emerging from North American research are:

- the time allocated for trainer-trainee interaction and making the best use of that time;
- methods of providing feedback;
- qualities of teachers rated as 'good' by trainees and their colleagues.

Methods for finding time for trainer-trainee interaction cover both time created within clinics by patient scheduling and the use of clinic experience for later case discussion and chart review. The papers in this area include many useful practical suggestions as well as evaluations of practice. In general, learning in ambulatory settings was found to give rise to discussion of a wider range of medical conditions, and greater attention to the medical interview and to social issues.

Issues relating to feedback included the findings that:

- supervisors changed their views quite often when they visited a patient after having heard the resident's report;

- counselling combined with detailed feedback on patient satisfaction scores improved the performance of residents with below average scores;
- the effectiveness of feedback tended to be reduced by "face-saving" strategies which ameliorated or even disguised the central message.

Qualities of good teachers inferred from rating studies can be grouped under the headings of Physician Role Model, Effective Supervisor, Dynamic Teacher and Supportive Person. There is a great deal of material in Section 4.5 which ought to be introduced into the training of clinical teachers.

Learning in inpatient settings is also researched in greater depth in North America. One gets the impression that American residents receive considerably more clinical teaching than their British counterparts, but there is no British data to enable a proper comparison. The variation in the amount of training received by British trainees is reported as considerable, raising issues of quality assurance and trainee entitlement. The two major constraining factors on learning by US residents were insufficient time and opportunity to learn, and low faculty involvement and commitment. Innovations receiving strong positive evaluations included adaptations of the Morning Report system to incorporate the teaching of Evidence-Based Medicine; and case reviews of patients whose diagnosis had changed while in hospital or within 6 months of leaving hospital.

Research on instructional thinking and decision-making by highly-rated clinical teachers is highly relevant to the training of clinical teachers ([Irb94a]). In addition to faculty development workshops on general teaching skills, case-based teaching is needed which can best be provided through departmental teaching improvement and mentoring skills.

9.6 Learning in Non-Clinical Settings

Research into postgraduate teaching and learning in non-clinical settings mostly comprises evaluations of a wide variety of teaching innovations, rich in ideas but not necessarily generalisable. In particular we would draw attention to improving the learning benefits of departmental conferences, developments in self-directed learning (more prominent for CME), skill-based courses in surgery, the use of GPs to teach primary care to house officers in Accident and Emergency departments, a system for teaching clinical examination comprising both seminar and ward-based components, and confirmation that various types of "lectures plus" teaching are more effective than lectures only. Several departures from the standard lecture format have been positively evaluated, as have variations on case-based departmental seminars. Thus the strong evidence that the effectiveness of off-the-job teaching is highly dependent on its links with related on-the-job teaching makes it unwise to evaluate off-the-job teaching on its own. This limits the applicability of some of the standard research on lecturing to undergraduates, for which research reviews are readily available.

9.7 Continuing Medical Education

Over the last decade there has been a gradual shift in focus from the provider-centred concept of Continuing Medical Education (CME) to the learner-centred concept of Con-

tinuing Professional Development (CPD). The recent Chief Medical Officer's Review of Continuing Professional Development and General Practice ([Cal98]) is an important indication of how government thinking has changed. The research evidence demonstrating that CME is only one of several contributors to physicians' learning and changes in their practice has been strong for some time, the other contributors being:

1. learning from other people through consultation and networking and
2. learning from publications and computerised information systems.

The relevant research falls into three main categories: research into how doctors learn, evaluation of CME interventions and research into innovation strategies using single or multiple interventions to achieve changes in specifically targeted areas of practice. All three of these interrelated areas of research have direct implications for practice.

Surveys of GPs, and also in a few cases consultants, have shown the importance for learning and changes in practice of a wide range of learning activities and sources of information. Moreover, they differ according to whether the changes involve treatment (including prescription), diagnosis and investigation, doctor-patient relationships, referral policy, health promotion or practice organisation. Models of physician learning distinguish between learning triggered by the problems raised by current individual patients and "learning projects" to acquire or improve proficiency in a targeted area of practice. The initiation of learning is dependent on significant background knowledge of what is out there to be learned to which CME conversations with other physicians, and reading contribute in ways which would not be revealed, for example, by evaluations of CME events. The importance of informal consultations with others and a reluctance to "cold call" experts suggest that facilitating social interaction among doctors and strengthening their networks should be a policy goal. These specifically medical models of physicians' learning, go into greater depth than general models of adult learning, though the latter are still confirmed by recent research. They enable more detailed discussions about professional learning, and especially lifelong learning, in which all physicians should now be prepared to participate.

Evaluations of CME courses have demonstrated the importance of including activities such as the observation and discussion of visual material and/or supervised practical work. Though it has confirmed that short courses of 1 day or less are rarely effective, no controlled studies have been reported which used length of course as a variable. This deficiency needs to be remedied because much time could be wasted trying to improve courses which are too short; and unrealistic expectations of the learning time required for certain goals are easily developed by busy learners and under-resourced providers — a form of collusion from which nobody benefits. Another important conclusion is that educational interventions on their own often fail to achieve changes in practice.

Research on innovation strategies points to the danger of focusing only on the development of competence. Competence has to be translated into performance and at this stage many dispositional and organisational factors come into play. Research on the implementation of guidelines, for example, indicates not only that the quality and utility of the guidelines themselves is important but also that both educational interventions (leading to understanding of their purpose and rationale) and administrative interventions (ranging from organisational changes to simple reminders) need to accompany the guidelines.

The discussion of recent developments in CPD reaches two conclusions. First, needs analysis is important for quality assurance purposes at three levels - the individual, the working group and the healthcare organisation (the last two are multi-professional). However, it should not be assumed that needs identified by audit, for example, will necessarily require an educational response. Second, following the advice of [FB98b], CME providers should adopt a coordinated approach to all three levels by facilitating self-directed learning, providing high quality individual and group education, and assisting healthcare organisations to develop and practise organisational learning.

9.8 Information Technology

Researchers foresee an increasing delivery of "processed data", e.g. about treatment options, at the point of care delivery, and the increasing dependence of practitioners on validated and trusted databases rather than the primary (journal) literature. They anticipate a coming together of evidence-based medicine and clinical information technology to provide the practitioner with whatever information is needed at the time and place where it is needed. Of course this raises the issues of how the practitioner is to be trained to access and to judge the trustworthiness of such information and how systems are to be designed to make information available in an effective manner that properly meets doctors' needs.

Decision support systems have had a mixed reception over the years, but seem to be growing in acceptance as just another tool in the doctor's armoury. A training issue here is the need to ensure that doctors understand how decision support systems frame the problem so that they can judge the quality of the advice that they offer.

In terms of the development of such systems, we may expect the development of linked databases rather than fragmented sources and better explanations from systems about the reasons for their decision advice.

A great many systems have been developed for various aspects of training, but most come up against the the consistently hard problem of having the system monitor, evaluate and react sensibly to the learner's attempts to master a skill or solve a problem. The training implication of this is that such systems may be excellent, but their use needs to be carefully integrated into the overall training programme — not least so that the human trainers provide what the computer-based system cannot provide in terms of monitoring and feedback.

9.9 Assessment and Revalidation

Hitherto, most of the research effort has focussed on the assessment of competence linked to certification decisions rather than the assessment of performance on-the-job. This emphasis is gradually changing as public demand for robust quality assurance grows. Revalidation is about to be developed with very close attention to research on performance assessment. The summative assessment of GPs is more competence-based as candidates provide their own sample of videorecorded patient consultations: one might argue that a random sample might be more appropriate if the training period were longer. Summative assessment for the award of the Certificate of Completion

of Specialist Training will be at least partly performance-based; and will need to be evaluated as it comes on line. Since Membership Examinations of the Royal Colleges are competence-based, the performance-based element at the end of Basic training is given relatively little attention. Thus the most critical certification issue arising from our review is the extent to which assessment regimes cover the full range of competence discussed in Chapter 3 and its translation into performance.

The other two purposes of assessment are (1) formative assessment to provide guidance to learners and/or those who supervise and support them, and (2) quality assurance and the improvement of practice. What research we have seen suggests that these issues deserve considerably more attention. With formative assessment questions have been raised about frequency, coverage, and reliability; and, if it is to properly serve its purpose, the manner in which formative assessment is integrated into training programmes to support the learning process will also need to be researched. Levels of supervision are often affected by factors other than the competence level of the trainee, and feedback may not be based on any systematic (though not necessarily formal) assessment.

The Canadian three-tier system for the monitoring and enhancement of physician performance is now well developed and familiar to those exploring revalidation in the UK. The need for good assessment practice linked with effective strategies for the improvement of practice (see Chapter 6) is critical, both for revalidation and for formative and summative assessments during postgraduate education.

9.10 Implications for Supervisory and Educational Roles

The practice of medicine involves dealing with complexity and uncertainty, often under considerable time pressure. Risk is inevitable, but can sometimes be reduced by the possession of, or ready access to, high levels of experience and expertise. One of the reasons for assigning house officers and registrars to a **firm** of consultants or a primary health care practice is to provide such access. There is also a risk in medical education. At some point a trainee has to conduct a procedure or make a decision for the first time; and taking such responsibility plays an important part in learning. Trainee professionals frequently report learning most intensively when 'on call' ([E...97, EACS98]) and it is through positive experiences of that kind that they develop their **confidence** to practice and to face new challenging situations. Medical culture places great importance on doctors' confidence being a reliable indicator of their competence; but recognises that practice also has to be situated within a framework of risk management and accountability to individual patients. How is this to be achieved when overconfidence is risky and lack of confidence has a negative impact on both patients and other health professionals?

Learning to assess and reassess one's own competence and its limits is a long and complex process, which becomes increasingly sophisticated as a doctor progresses through postgraduate education. Its reliability significantly depends on access to good supervision and feedback. Feedback which contributes to a trainee doctor's self-assessment may come from patient outcomes, informal discussions with other doctors or health professionals, periodic appraisals or meetings for signing the trainee's logbook. Informal feedback on the wards tends to be spontaneous and incidental, i.e. not the result of a reflective judgement: it is mostly received from more senior trainees rather than consultants, and more likely to be negative than positive. Formal feedback appears to vary

considerably in quantity, quality and breadth of coverage. Even the best designed log-books focus on competence rather than performance. Techniques such as chart review which emphasise performance are rarely used in Britain outside General Practice. As reported in Chapter 4, there is sufficient cause for concern to suggest that research into the practice and effectiveness of supervision and feedback during postgraduate medical education is urgently needed. This should include the implicit delegation of certain supervisory and educational responsibilities to senior trainees. Should it be formalised, as in the US role of Chief Resident? Should it be accompanied by training? Which responsibilities could or should be delegated, and which should not?

The other key educational role is that of facilitating learning from clinical experience. For this there is a continuum of possible clinical educators from those who bring clinical and educational expertise to lectures, seminars or workshops for basic, higher specialist, or general practice trainees to more senior trainees who work very closely with their less experiences colleagues in the wards and are best positioned to take up 'live' learning opportunities as they happen. In between are the consultants in their firm who have greater expertise and normally are reasonably accessible for discussing significant cases linked to their specialisms. In practice, specialist registrars will expect to learn most from the consultants, SHOs from registrars and PRHOs from SHOs. But these arrangements are informal. It could be a junior SHO who judges when a PRHO is working within their level of competence, and who might occasionally think about appropriate learning opportunities for which they have no formal responsibility; and junior SHOs have little preparation or experience for this role. Moreover, changes in rotation systems to reduce junior doctors' hours have resulted in many hospitals in working groups of house officers and registrars whose membership changes as frequently as every 6 weeks. The transient nature of these groups leaves little time for them to get to know each other well enough for mutual learning to reach an optimal level.

Chapter 4 reviewed a wide range of naturally occurring and deliberately created learning opportunities in clinical settings, for most of which there is good evidence of positive learning outcomes. The principal problems are those of creating greater awareness of these opportunities and developing clinical educators with the disposition to use them in developing educational practices appropriate to their circumstances. [Irb94a, Irb95] provides specific advice on the flexible training of clinical teachers. The main constraint on such developments is the belief that they are impossible to implement while current service pressures on doctors' time persist. This management issue, which is not confined to medical education, requires more attention. Few Trusts have reliable mechanisms for incorporating clinical teaching into their organisation of professional time, though some may have formally agreed to do so. Nor are there any internal audit mechanisms at local level for monitoring and periodically evaluating a Trust's programme of professional education. Clinical Tutors have neither the time nor the authority to undertake such duties.

The role of mentoring has recently been discussed by [Sta99] and reviewed by [Bli99]. There is little evidence about its use in medicine, unless one includes the peer-tutoring experiments among GPs reported in Chapter 6. Its use in business is frequently advocated and sometimes practised ([EACS99]), but has given rise to little research. The term is used in teacher education with a rather different and, in our view, less authentic meaning - that of a placement supervisor with a practice teaching role, who is often also an assessor. Given the difficulties discussed above of finding sufficient time for

supervision, giving feedback and clinical teaching for postgraduate medical trainees, the introduction of mentoring as an additional role and obligation might not justify a high priority. However, mentoring might be particularly well suited to the support of doctors during the first few years after completion of postgraduate training. Both consultants and GPs could benefit from such support as they grow into their new roles and responsibilities, learn to work with new colleagues and to contribute appropriately to their Trust or General Practice, and take greater responsibility for organising their own lifelong learning.

Bibliography

[ACJ91] Carl C. Awh, Howard P. Cupples, and Jonathan C. Javitt. Improved detection and referral of patients with diabetic retinopathy by primary care physicians. effectiveness of education. *Archives of Internal Medicine*, 151(7):1405–1408, 1991.

[Ada87] Jack A. Adams. Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. *Psychological Bulletin*, 101(1):41–74, 1987.

[AL98] Roger Azevedo and Susanne P. Lajoie. The cognitive basis for the design of a mammography interpretation tutor. *International Journal of Artificial Intelligence in Education*, 9:32–44, 1998.

[ALD...97] Roger Azevedo, Susanne P. Lajoie, Monique Desaulniers, David M. Fleiszer, and Patrice M. Bret. Radtutor: The theoretical and empirical basis for the design of a mammography interpretation tutor. In *Proceedings of World Conference on Artificial Intelligence and Education, AIED'95*, Kobe, Japan, 19-22 Aug 1997 1997.

[Alg98] P. C. Alguire. A review of journal clubs in postgraduate medical education. *Journal of General Internal Medicine*, 13(5):347–353, 1998.

[AM93] M. A. Albanese and S. Mitchell. Problem-based learning: a review of the literature on its outcomes and implementation issues. [published erratum appears in acad med 1993 aug;68(8):615] [see comments]. *Academic Medicine*, 68(1):52–81, 1993. United States.

[Ame89] American College of Physicians. Working-conditions and supervision for residents in internal medicine programs - recommendations. *Annals of Internal Medicine*, 110(8):657–663, 1989.

[And86] W. T. Anderson. The apostolic function of the dentist. In Sue Fisher and Alexandra Dundas Todd, editors, *Discourse and institutional authority : medicine, education and law*, number 19 in Advances in discourse processes. Ablex, Norwood, N.J., 1986.

[AOR97] L. A. Allery, P. A. Owen, and M. R. Robling. Why general practitioners and consultants change their clinical practice: a critical incident study [see comments]. *British Medical Journal*, 314(7084):870–874, 1997. England.

[APP93] J. F. Arocha, V. L. Patel, and Y. C. Patel. Hypothesis generation and the coordination of theory and evidence in novice diagnostic reasoning. *Medical Decision Making*, 13(3):198–211, 1993.

[ARJ96] D. Armstrong, H. Reyburn, and R. Jones. A study of general practitioners' reasons for changing their prescribing behaviour. *British Medical Journal*, 312(7036):949–952, 1996. England.

[ASM...89] B. O. Anderson, J. H. Sun, E. E. Moore, L. L. Thompson, A. H. Harkin, and E. J. Bartle. The development and evaluation of a clinical test of surgical resident proficiency. *Surgery*, 106(2):347–3, 1989. United States.

[ASST93] A. Al-Shehri, I. Stanley, and P. Thomas. Continuing education for general practice. Part 2. Systematic learning from experience. *British Journal of General Practice*, 43:249–253, 1993.

[ATS...99] R. J. Aitken, M. R. Thompson, J. A. Smith, A. G. Radcliffe, J. D. Stamatakis, and R. J. Steele. Training in large bowel cancer surgery: observations from three prospective regional united kingdom audits [see comments]. *British Medical Journal*, 318(7185):702–703, 1999. England.

[B...94] A. Barker et al. A comparative study of postgraduate medical education in north east thames region. *Postgraduate Medical Journal*, 70:722–727, 1994.

[Bah98] Jamie Bahrami. Yorkshire personal education plan: report of a workshop. In R. Spencer-Jones, editor, *Self-directed learning for GPs*, pages 520–521. 1998. Supplement to: Education for General Practice; v9 no.4 (supp.); s497-s556.

[Bai76] D. J. Bain. Doctor-patient communication in general practice consultations. *Medical Education*, 10:125–131, 1976.

[Bai77] D. J. Bain. Patient knowledge and the content of the consultation in general practice. *Medical Education*, 11:347–350, 1977.

[Bal57] Michael Balint. *The doctor, his patient and the illness*. Pitman, London, 1957.

[Bar87] R. W. Barnes. Surgical handicraft - teaching and learning surgical skills. *American Journal of Surgery*, 153(5):422–427, 1987.

[Bar93] H. S. Barrows. An overview of the uses of standardized patients for teaching and evaluating clinical skills. aamc. *Academic Medicine*, 68(6):443–451, 1993. discussion451-3.

[BBG97] K. E. Blackwell, M. T. Brown, and D. Gonzalez. Overcoming the learning curve in microvascular head and neck reconstruction. *Arch Otolaryngol Head Neck Surg*, 123(12):1332–1335, 1997.

[BBIS98] Georges Bordage, Jeffrey H. Burack, David M. Irby, and Frank T. Stritter. Education in ambulatory settings: Developing valid measures of educational outcomes, and other research priorities. *Academic Medicine*, 73(7):743–750, 1998. United States.

[BBM98] G. A. Bunch, J. Bahrami, and R. Macdonald. Basic surgical training: How good is it? *Annals of the Royal College of Surgeons of England*, 80:219–222, 1998.

[Bec79] M. H. Becker. Understanding patient compliance: the contributions of attitudes and other psychosocial factors. In Stuart J. Cohen, editor, *New Directions on Patient Compliance*, pages 171–171. Lexington, Mass, Lexington, 1979.

[BJ97] A. J. Byrne and J. G. Jones. Responses to simulated anaesthetic emergencies by anaesthetists with different durations of clinical experience [see comments]. *British Journal of Anaesthesia*, 78(5):553–556, 1997. Comment in: Br J Anaesth 1997 Nov;79(5):689; discussion 690-1; Comment in: Br J Anaesth 1997 Nov;79(5):689-90; discussion 690-1.

[BL76] Patrick Sarsfield Byrne and Barrie E. L. Long. *Doctors talking to patients : a study of the verbal behaviour of general practitioners consulting in their surgeries.* HMSO, London, 1976.

[Bli99] John Bligh. Mentoring: an invisible support network. *Medical Education*, 33:2–3, 1999.

[BMF96] N. C. Boreham, G. E. Mawer, and R. W. Foster. Medical diagnosis from circumstantial evidence. *Le Travail Humain*, 59(1):69–85, 1996.

[BN97] B. G. Brennan and G. R. Norman. Use of encounter cards for evaluation of residents in obstetrics. *Academic Medicine*, 72(10, 1):S43–S44, 1997.

[BNC93] C. J. Babcook, G. R. Norman, and C. L. Coblentz. Effect of clinical history on the interpretation of chest radiographs in childhood bronchiolitis. *Investigative Radiology*, 28(3):214–217, 1993.

[BNC...95] F. P. Bryce, R. G. Neville, I. K. Crombie, R. A. Clark, and P. McKenzie. Controlled trial of an audit facilitator in diagnosis and treatment of childhood asthma in general practice [see comments]. *British Medical Journal*, 310(6983):838–842, 1995. England.

[Bog97] M. S. Bognor. Naturalistic decision making in health care. In Caroline E. Zsambok and Gary Klein, editors, *Naturalistic Decision Making (Expertise : Research and Applications)*, pages 61–70. Lawrence Erlbaum Associates, Mahwah, N.J., 1997.

[Bor89] N. C. Boreham. Modelling medical decision-making under uncertainty. *British Journal of Educational Psychology*, 59:187–199, 1989.

[Bor91] G. Bordage. Theory practice theory - a healthy cycle in medical-education -foreword. *Academic Medicine*, 66(9, S):SR7–SR7, 1991.

[Bor94] N. C. Boreham. The dangerous practice of thinking. *Medical Education*, 28(3):172–179, 1994.

[BOSR96] L. P. Burke, N. A. Osborn, J. E. Smith, and A. P. Reid. Learning fibreoptic skills in ear, nose and throat clinics. *Anaesthesia*, 51(1):81–83, 1996.

[Bou99] P. Bourdillon. Dutch system of peer review is different and effective [letter]. *British Medical Journal*, 318(7191):1143–1143, 1999.

[BP87] G. Bordage and G. Page. An alternate approach to pmps: the key features concept. In Ian R. Hart and Ronald M. Harden, editors, *Further developments in assessing clinical competence*. Can-Heal Publications, Montreal, 1987. 1987.

[BRP91] K. D. Bertakis, D. Roter, and S. M. Putnam. The relationship of physician medical interview style to patient satisfaction [see comments]. *Journal of Family Practice*, 32(2):175–181, 1991. Comment in: J Fam Pract 1991 Feb;32(2):135-6.

[BS85] J. M. Bensing and E. M. Sluijs. Evaluation of an interview training course for general practitioners. *Social Science and Medicine*, 20(7):737–744, 1985.

[BSJ98] A. J. Byrne, A. J. Sellen, and J. G. Jones. Errors on anaesthetic record charts as a measure of anaesthetic performance during simulated critical incidents. *British Journal of Anaesthesia*, 80(1):58–62, 1998.

[BTR90] P. G. Bain, W. L. Tek, and T. J. Rowe. Workload of pre-registration house officers. *Medical Journal*, 300:1463–1463, 1990.

[Bul93] R. Bulger. On the Oxford perinatal care model and medical education. *International Journal of Technology Assessment in Health Care*, 9(3):51–59, 1993.

[Bur98] Jonathan Burton. Multipractice, self-directed learning groups in north thames east region. In R. Spencer-Jones, editor, *Self-directed learning for GPs*, pages 512–516. 1998. Supplement to: Education for General Practice; v9 no.4 (supp.); s497-s556.

[Cal98] Kenneth C. Calman. *A review of continuing professional development in general practice*. Department of Health, Leeds, England, 14/5/98 edition, 1998.

[CBC98] Rowland W. Chang, Georges Bordage, and Karen J. Connell. The importance of early problem representation during case presentations. *Academic Medicine*, 73(10 suppl):s109–s111, 1998.

[CD91] K. C. Calman and M. Donaldson. The pre-registration house officer year: a critical incident study. *Medical Education*, 25(1):51–59, 1991. England.

[CDFM95] P. A. Carney, A. J. Dietrich, Jr. Freeman, D. H., and L. A. Mott. A standardized-patient assessment of a continuing medical education program to improve physicians' cancer-control clinical skills. *Academic Medicine*, 70(1):52–58, 1995. United States.

[CGdJ...94] V. Chopra, B. J. Gesink, J. de Jong, J. G. Bovill, J. Spierdijk, and R. Brand. Does training on an anaesthesia simulator lead to improvement in performance? *British Journal of Anaesthesia*, 73(3):293–297, 1994.

[CH95] T. Carney and C. Helliwell. Effect of structured postgraduate medical-education on the care of patients with diabetes. *British Journal of General Practice*, 45(392):149–151, 1995.

[CHM95] L. M. Campbell, J. G. Howie, and T. S. Murray. Use of videotaped consultations in summative assessment of trainees in general practice. *British Journal of General Practice*, 45(392):137–141, 1995.

[CJ99] P. Cantillon and R. Jones. Does continuing medical education in general practice make a difference? [in process citation]. *British Medical Journal*, 318(7193):1276–1279, 1999.

[CLLB86] D. W. Cope, L. S. Linn, B. D. Leake, and P. A. Barrett. Modification of residents' behavior by preceptor feedback of patient satisfaction. *Journal of General Internal Medicine*, 1(6):394–398, 1986.

[CMHF97] M. Challis, N. J. Mathers, A. C. Howe, and N. J. Field. Portfolio-based learning: Continuing medical education for general practitioners - a mid-point evaluation. *Medical Education*, 31(1):22–26, 1997.

[Col99] C. Collins. Surgical training, supervision, and service [editorial; comment]. *British Medical Journal*, 318(7185):682–683, 1999. England.

[CP87] L. D. Coughlin and V. L. Patel. Processing of critical information by physicians and medical-students. *Journal of Medical Education*, 62(10):818–828, 1987.

[CP98] K. N. Cytryn and V. L. Patel. Reasoning about diabetes and its relationship to the use of telecommunication technology by patients and physicians. *International Journal of Medical Informatics*, 51(2-3):137–151, 1998.

[CRM...96] D. M. Chapman, K. J. Rhee, J. A. Marx, B. Honigman, E. A. Panacek, D. Martinez, B. T. Brofeldt, and S. H. Cavanaugh. Open thoracotomy procedural competency: validity study of teaching and assessment modalities [see comments]. *Annals of Emergency Medicine*, 28(6):641–647, 1996. Comment in: Ann Emerg Med 1997 Jul;30(1):117-8.

[CRN96] Eugene J. F. M. Custers, Glenn Regehr, and Geoffrey R. Norman. Mental representations of medical diagnostic knowledge: A review. *Academic Medicine*, 71(10 suppl, S):S55–S61, 1996.

[CRV...96] D. H. Cordes, Joseph L. Rea, Anthony F. Vuturo, Evan W. Kligman, and Dorothy Foster Rea. A survey of residency management training: general preventive medicine graduates. *American Journal of Preventive Medicine*, 12(3):172–176, 1996.

[CTN...99] K. C. Calman, J. G. Temple, R. Naysmith, R. G. Cairncross, and S. J. Bennett. Reforming higher specialist training in the united kingdom–a step along the continuum of medical education. *Medical Education*, 33(1):28–33, 1999. England.

[Cun95] William Cunningham. Interdisciplinary practice educational meetings - continuing education for the working primary health care team. *Education for General Practice*, 6:41–48, 1995.

[CYGR...88] C. Chan-Yan, J. H. Gillies, J. Ruedy, J. S. Montaner, and S. A. Marshall. Clinical skills of medical residents: a review of physical examination. *Canadian Medical Assocition Journal*, 139(7):629–632, 1988.

[Dau95] W. D. Dauphinee. Assessing clinical performance. where do we stand and what might we expect? [editorial; comment]. *Journal of the American Medical Association*, 274(9, 6 Sept 1995):741–743, 1995. Comment on: JAMA 1995 Sep 6;274(9):735-40.

[DB91] Sue Dowling and Sue Barrett. *Doctors in the making : the experience of the pre-registration year*. University of Bristol, Department of Epidemiology and Public Health Medicine, Bristol, England, 1991.

[DBSF98] A. M. Derossis, J. Bothwell, H. H. Sigman, and G. M. Fried. The effect of practice on performance in a laparoscopic simulator. *Surgical Endoscopy-Ultrasound and Interventional Techniques*, 12(9):1117–1120, 1998.

[DDB...83] D. A. Davis, T. Delmore, A. M. Bryans, J. I. Williams, E. Dunn, J. Krauser, D. J. Scott, and A. Heron. Continuing medical education in ontario. *Annals of the Royal College of Physicians and Surgeons of Canada*, 16:136–141, 1983.

[DDS...97] D. A. DaRosa, G. L. Dunnington, J. Stearns, G. Ferenchick, J. L. Bowen, and D. E. Simpson. Ambulatory teaching "lite": less clinic time, more educationally fulfilling [see comments]. *Academic Medicine*, 72(5):358–361, 1997. Comment in: Acad Med 1997 May;72(5):357.

[DDVG97] Jeremy Dale, Mark Davies, Kishor Vassant, and Edward Glucksman. Primary care consultation skills training: implementing a new training programme for senior house officers in an accident and emergency department. *Medical Education*, 31(4):243–249, 1997.

[DE88] Jack Dowie and Arthur Elstein. *Professional judgement - a reader in clinical decision making*. Cambridge University Press, Cambridge, England, 1991 reprint edition, 1988.

[DFA...98] A. M. Derossis, G. M. Fried, M. Abrahamowicz, H. H. Sigman, J. S. Barkun, and J. L. Meakins. Development of a model for training and evaluation of laparoscopic skills. *American Journal of Surgery*, 175(6):482–487, 1998.

[DFJ...94] Dale Dauphinee, Wesley Fabb, Brian Jolly, Donald Langsley, Stephen Wealthall, and Peter Procopis. Determining the content of certifying

examinations. In David Newble, Brian Jolly, and Richard Wakeford, editors, *The certification and recertification of doctors: issues in the assessment of clinical competence*, pages 92–104. Cambridge University Press, Cambridge, England, 1994.

[DJ95] A. C. Dumay and G. J. Jense. Endoscopic surgery simulation in a virtual environment. *Computers in Biology and Medicine*, 25(2):139–148, 1995.

[DLG...91] J. Derry, M. Lawrence, K. Griew, J. Anderson, J. Humphreys, and K. S. Pandher. Auditing audits: the method of Oxfordshire medical audit advisory group [see comments]. *British Medical Journal*, 303(6812):1247–1249, 1991. Comment in: BMJ 1992 Feb 22;304(6825):509.

[DLM94] David A. Davis, Elizabeth Lindsay, and Paul Mazmanian. The effectiveness of CME interventions. In David A. Davis and Robert D. Fox, editors, *The physician as learner - linking research to practice*, pages 241–278. American Medical Association, 1994.

[DOK...92] A. J. Dietrich, G. T. O'Connor, A. Keller, P. A. Carney, D. Levy, and F. S. Whaley. Cancer: improving early detection and prevention. a community practice randomised trial [see comments]. *British Medical Journal*, 304(6828):687–691, 1992. Comment in: BMJ 1992 Mar 14;304(6828):652-3; Comment in: BMJ 1992 May 2;304(6835):1177.

[Don94] L. J. Donaldson. Doctors with problems in an NHS workforce [see comments]. *British Medical Journal*, 308(6939):1277–1282, 1994. England.

[Dow93] Jack Dowie. Clinical decision analysis: background and introduction. In Huw Llewelyn and Anthony Hopkins, editors, *Analysing how we reach clinical decisions*, pages 7–26. Royal College of Physicians of London, London, 1993.

[DTOH95] David A. Davis, Mary Ann Thompson, Andrew D. Oxman, and Brian Haynes. Changing physician performance a systematic review of the effect of continuing medical education strategies. *Journal of the American Medical Association*, 274(9):700–705, 1995.

[DWHO96] M. J. Donnelly, M. A. Walsh, S. Hone, and P. O'Sullivan. Examining the ear: clinical teaching. *Medical Education*, 30(4):299–302, 1996.

[DWW94] Martin Drage, Richard Wakeford, and Amanda Wharton. What do general practitioners think changes their clinical behaviour? *Education for General Practice*, 5:48–53, 1994.

[E...97] Michael Eraut et al. *Improving the training of pre-registration house officers: report to BPMF*. University of Sussex Institute of Education, Brighton, England, 1997.

[EABW95] M. Eraut, J. Alderton, A. Boylan, and A. Wraight. *Learning to Use Scientific Knowledge in Education and Practice Settings*. English National Board for Nursing, Midwifery and Health Visiting, London, 1995.

[EACS98] M. Eraut, J. Alderton, G. Cole, and P Senker. Development of knowledge and skills in employment. Research Report 5, University of Sussex Institute of Education, 1998.

[EACS99] Michael Eraut, J. Alderton, G. Cole, and P. Senker. The impact of the manager on learning in the workplace. In Frank Coffield, editor, *Speaking truth to power: research and policy in the workplace*. 1999.

[EFC97] R. B. Elson, J. G. Faughnan, and D. P. Connelly. An industrial process view of information delivery to support clinical decision making: Implications for systems design and process measures. *Journal of the American Medical Informatics Association*, 4(4):266–278, 1997.

[EFW...96] A. S. Elstein, C. P. Friedman, F. M. Wolf, G. Murphy, J. Miller, P. Fine, P. Heckerling, T. Miller, J. Sisson, S. Barlas, K. Biolsi, M. Ng, X. Mei, T. Franz, and A. Capitano. Effects of a decision support system on the diagnostic accuracy of users: A preliminary report. *Journal of the American Medical Informatics Association*, 3(6):422–428, 1996.

[EKS...87] B. J. Evans, F. D. Kiellerup, R. O. Stanley, G. D. Burrows, and B. Sweet. A communication skills programme for increasing patients' satisfaction with general practice consultations. *British Journal of Medical Psychology*, 60(4):373–378, 1987.

[EM96] M. Eliastam and T. Mizrahi. Quality improvement, housestaff, and the role of chief residents. *Academic Medicine*, 71(6):670–674, 1996.

[EMC98] Michael Eraut, Louise Morley, and Gerald Cole. *Standards and vocational qualifications in continuing professional development (CPD)*. Qualifications and Curriculum Authority, London, England, 1998.

[EP89] David A. Evans and Vimla L. Patel. *Cognitive science in medicine - biomedical modelling*. M.I.T. Press, Cambridge, Massachusetts, 1989.

[EPE95] J. Ende, A. Pomerantz, and F. Erickson. Preceptors' strategies for correcting residents in an ambulatory care medicine setting: a qualitative analysis. *Academic Medicine*, 70(3):224–229, 1995.

[Era97] Michael Eraut. Perspectives on defining 'The Learning Society'. *Journal of Education Policy*, 12(6):551–558, 1997.

[Era98] Michael Eraut. Concepts of competence. *Journal of Interprofessional Care*, 12(2):127–139, 1998.

[ES97] S. Ebrahim and G. D. Smith. Systematic review of randomised controlled trials of multiple risk factor interventions for preventing coronary heart disease [see comments]. *British Medical Journal*, 314(7095):1666–1674, 1997. England.

[Evi92] Evidence-Based Medicine Working Group: McMaster University. Evidence-based medicine. A new approach to teaching the practice of medicine. *Journal of American Medical Association*, 268(17):2420–2425, 1992. Comment in: JAMA 1993 Mar 10;269(10):1253; discussion 1254.

[FB98a] S. L. Fowell and J. G. Bligh. Recent developments in assessing medical students. *Postgraduate Medical Journal*, 74(867):18–24, 1998.

[FB98b] R. D. Fox and N. L. Bennett. Learning and change: implications for continuing medical education. *British Medical Journal*, 316(7129):466–468, 1998. England.

[FC98] Della Fish and Colin Coles. *Developing Professional Judgement in Health Care : learning through the critical appreciation of practice.* Butterworth-Heinemann, Oxford, 1998.

[FCSDS92] Paul J. Feltovich, Richard L. Coulson, Rand J. Spiro, and Beth K. Dawson-Saunders. Knowledge application and transfer for complex tasks in ill-structured domains: implications for instruction and testing in biomedicine. In David A. Evans and Vimla L. Patel, editors, *Advanced Models of Cognition for Medical Training and Practice*, pages 213–244. Springer-Verlag, Berlin Heidelberg, 1992. Proceedings of the NATO Advanced Research Workshop on Advanced Models of Cognition for Medical Training and Practice held at Il Ciocco, Barga, Italy June 19-22 1991.

[FDW94] Robert D. Fox, David A. Davis, and Dennis Wentz. The case for research on continuing medical education. In David A. Davis and Robert D. Fox, editors, *The physician as learner - linking research to practice*, pages 15–26. American Medical Association, 1994.

[FE98] John G. Faughnan and Robert Elson. Information technology and the clinical curriculum: Some predictions and their implications for the class of 2003. *Academic Medicine*, 73(7):766–769, 1998.

[FF88] Charles Fletcher and Paul Freeling. *Talking and listening to patients: a modern approach.* Nuffield Provincial Hospital Trust, London, 1988.

[FFG87] E. Fullard, G. Fowler, and M. Gray. Promoting prevention in primary care: controlled trial of low technology, low cost approach. *British Medical Journal*, 294(6579):1080–1082, 1987. England.

[FGH...95] G. Feder, C. Griffiths, C. Highton, S. Eldridge, M. Spence, and L. Southgate. Do clinical guidelines introduced with practice based education improve care of asthmatic and diabetic patients? A randomised controlled trial in general practices in east london. *British Medical Journal*, 311(7018):1473–1478, 1995. England.

[FMM94] R. C. Fraser, R. K. McKinley, and H. Mulholland. Consultation competence in general practice: testing the reliability of the Leicester assessment package. *British Journal of General Practice*, 44:293–296, 1994.

[FMP89] R. D. Fox, Paul Mazmanian, and R. W. Putnam. *Change and learning in the lives of physicians.* Praeger Publishing, New York, NY, 1989.

[For94] K. A. Forde. Minimal access surgery–which path to competence? [editorial; comment]. *Surgical Endoscopy-Ultrasound and Interventional Techniques*, 8(9):1047–1048, 1994. Germany. Comment on: Surg Endosc 1994 Sep;8(9):1137-42.

[Fri96] Richard Bruce Friedman. Top ten reasons the world wide web may fail to change medical education. *Academic Medicine*, 71(9):979–981, 1996.

[FSB...97] G. Ferenchick, D. Simpson, J. Blackman, D. DaRosa, and G. Dunnington. Strategies for efficient and effective teaching in the ambulatory care setting. *Academic Medicine*, 72(4):277–280, 1997.

[FSK95] I. G. Finlay, N. C. H. Stott, and P. Kinnersley. The assessment of communication skills in palliative medicine: a comparison of the scores of examiners and simulated patients. *Medical Education*, 29:424–429, 1995.

[FWT93] William F. Jr Fallon, Robert L. Wears, and III Tepas, Joseph J. Resident supervision in the operating room: does this impact on outcome. *Journal of Trauma*, 35(4):556–561, 1993.

[Gab92] David M. Gaba. Dynamic decision-making in anesthesiology: cognitive models and training approaches. In David A. Evans and Vimla L. Patel, editors, *Advanced Models of Cognition for Medical Training and Practice*, pages 123–148. Springer-Verlag, Berlin Heidelberg, 1992. Proceedings of the NATO Advanced Research Workshop on Advanced Models of Cognition for Medical Training and Practice held at Il Ciocco, Barga, Italy June 19-22 1991.

[GBI91] Timothy F. Gabryel, Marcia A. Brierley, and Joseph F. Izzo. How completely internal medicine residents at a new york state teaching hospital met the expectations of a credentialling protocol. *Academic Medicine*, 66(10):613–615, 1991.

[GDA...93] Jonathan H. Gillard, Thomas H.S. Dent, Emma J. Aarons, Penelope J. Smyth-Pigott, and Michael W.N. Nicholls. Prergistration house officers in eight English regions: survey of qulaity of training. *British Medical Journal*, 307:1180–1184, 1993.

[GDT...98] R. Grol, J. Dalhuijsen, S. Thomas, C. Veld, G. Rutten, and H. Mokkink. Attributes of clinical guidelines that influence use of guidelines in general practice: observational study. *British Medical Journal*, 317(7162):858–861, 1998. England.

[GG93] V. M. Gennis and M. A. Gennis. Supervision in the outpatient clinic: effects on teaching and patient care [see comments]. *Journal of General Internal Medicine*, 8(7):378–380, 1993. Comment in: J Gen Intern Med 1994 Feb;9(2):116-7.

[Gle97] Fergus Gleeson. AMEE medical education guide no. 9. Assessment of clinical competence using the objective structured long examination record (osler). *Medical Teacher*, 19:7–14, 1997.

[GM88] J. Grant and P. Marsden. Primary knowledge, medical-education and consultant expertise. *Medical Education*, 22(3):173–179, 1988.

[GM92] J. Grant and P. Marsden. *Training senior house officers by service-based learning.* Joint Centre for Education in Medicine, London, 1992.

[GM97] Charles Galasko and Colin MacKay. Unsupervised surgical training - logbooks are essential for assessing progress. *British Medical Journal*, 315(15 nov), 1997.

[GMHL...89] R. Grol, H. Mokkink, A. Helsper-Lucas, V. Tielens, and J. Bulte. Effects of the vocational training of general practice consultation skills and medical performance. *Medical Education*, 23(6):512–521, 1989.

[GMK89] J. Grant, P. Marsden, and R. C. King. Senior house officers and their training. i. personal characteristics and professional circumstances. *British Medical Journal*, 299(6710):1263–1265, 1989. England.

[GMS94] J. Gear, A. McIntosh, and G. Squires. *Informal learning in the professions.* Department of Adult Learning, University of Hull, 1994.

[GNW...93] G. H. Guyatt, J. Nishikawa, A. Willan, W. McIlroy, D. Cook, J. Gibson, A. Kerigan, and A. Neville. A measurement process for evaluating clinical teachers in internal medicine [see comments]. *Canadian Medical Association Journal*, 149(8):1097–1102, 1993. Comment in: Can Med Assoc J 1993 Oct 15;149(8):1085; Comment in: Can Med Assoc J 1994 Oct 1;151(7):908-9.

[Gor84] M. D. Gordon. Helping general practitioners to keep up with the literature: evaluation of an RCGP initiative. *Medical Education*, 18(3):174–177, 1984. England.

[GPW82] R. H. Geertsma, Jr. Parker, R. C., and S. K. Whitbourne. How physicians view the process of change in their practice behavior. *Journal of Medical Education*, 57:752–761, 1982.

[GR93] J. M. Grimshaw and I. T. Russell. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations [see comments]. *Lancet*, 342:1317–1322, 1993.

[Gra89] Janet Grant. Clinical decision making: Rational principles, clinical intuition or clinical thinking. In John I. Balla, Margaret Gibson, and Anne M. Chang, editors, *Learning in Medical School, A Model for the Clinical Professions*, pages 81–102. 1989.

[Gra98] R. Gray. *The trainer/registrar tutorial in medical general practice postgraduate education.* MA dissertation, University of Sussex, 1998.

[GRHW97] Charles H. I. I. I. Griffith, Eugene C. Rich, Steven D. Hillson, and John F. Wilson. Internal medicine residency training and outcomes [see comments]. *Journal of General Internal Medicine*, 12(6):390–396, 1997. Comment in: J Gen Intern Med 1997 Jun;12(6):399-400.

[Gru97] Larry D. Gruppen. Implications of cognitive research for ambulatory care education. implications of cognitive research for ambulatory care education. *Academic Medicine*, 72(2):117–120, 1997.

[GSFHR95] Afaf Girgis, Rob W. Sanson-Fisher, Chris Howe, and Bruce Raffan. A skin cancer training programme: evaluation of a postgraduate training for family doctors. *Medical Education*, 29(5):364–371, 1995.

[GSL97] Michael Goldacre, Sarah Stear, and Trevor Lambert. The pre-registration year. the trainees' experience. *Medical Education*, 31(Suppl 1):57–60, 1997.

[GTW...94] Richard A. Garibaldi, Marie C. Trontell, Herbert Waxman, John H. Holbrook, Theresa D. Kanya, Shahram Koshbin, John Thompson, Maribeth Casey, Raja G. Subhiya, and Frank Davidoff. The in-training examination in internal medicine. *Annals of Internal Medicine*, 121:117–123, 1994.

[GUTW98] Linda Gask, Tim Usherwood, Hannah Thompson, and Bill Williams. Evaluation of a training package in the assessment and management of depression in primary care. *Medical Education*, 32:190–198, 1998.

[HAB95] Jean-Michel M. Hoc, Ren Amalberti, and Nicholas Boreham. Human operator expertise in diagnosis, decision-making, and time management. In Jean-Michel M. Hoc, Erik Hollnagel, and Pietro C. Cacciabue, editors, *Expertise and technology:Cognition & Human-Computer Cooperation*, pages 19–42. Lawrence Erlbaum Associates, Hove, 1995.

[HB98] Arthur Hibble and Robert Berrington. Personal professional learning plans - an evaluation of mentoring and co-tutoring in general practice. *Education for General Practice*, 9:216–221, 1998.

[HBC95] J. Heppell, G. Beauchamp, and A. Chollet. Ten-year experience with a basic technical skills and perioperative management workshop for first-year residents [see comments]. *Canadian Journal of Surgery*, 38(1):27–32, 1995. Comment in: Can J Surg 1995 Feb;38(1):8-9.

[HC98] A. C. Hartmann and Jr. Cruz, P. D. Interactive mechanisms for teaching dermatology to medical students [see comments]. *Archives of Dermatology*, 134(6):725–728, 1998. Comment in: Arch Dermatol 1998 Jun;134(6):731-2; Arch Dermatol 1998 Jun;134(6):725-8.

[HCBN96] B. Hodges, R. Cohen, A. Bienenstock, and G. Norman. Evaluating communication skills in the objective structured clinical examination format: reliability and generalizability. *Medical Education*, 30:38–43, 1996.

[HCG98] J. H. Hindmarsh, G. D. Coster, and C. Gilbert. Are vocationally trained general practitioners better GPs? a review of research designs and outcomes. *Medical Education*, 32(3):244–254, 1998.

[Hel81] C. G. Helman. Disease versus illness in general practice. *Journal of the Royal College of General Practitioners*, 27:277–279, 1981.

[Hew91] Mariana G. Hewson. Reflection in clinical teaching: an analysis of reflection-on-action and its implications for staffing residents. *Medical Teacher*, 13(3):227–231, 1991.

[HHM...75] J. R. Hampton, M. J. Harrison, J. R. Mitchell, J. S. Prichard, and C. Seymour. Relative contributions of history-taking, physical examination, and laboratory investigation to diagnosis and management of medical outpatients. *British Medical Journal*, 2:486–489, 1975.

[HL92] R. M. Harden and J. M. Laidlaw. Effective continuing-education — the crisis criteria. *Medical Education*, 26(5):408–422, 1992.

[HLKM96a] R. M. Harden, J. M. Laidlaw, J. S. Ker, and H. E. Mitchell. Amee medical-education guide no-7 — task-based learning — an educational strategy for undergraduate, postgraduate and continuing medical-education .1. *Medical Teacher*, 18(1):7–13, 1996.

[HLKM96b] R. M. Harden, J. M. Laidlaw, J. S. Ker, and H. E. Mitchell. Amee medical-education guide no-7 — task-based learning — an educational strategy for undergraduate, postgraduate and continuing medical-education .2. *Medical Teacher*, 18(2):91–98, 1996.

[HNB96] Rose A. Hatala, Geoffrey R. Norman, and Lee R. Brooks. The effect of clinical history on physicians' ECG interpretation skills. *Academic Medicine*, 71(10, S):S68–S70, 1996.

[HRHM98] Brian Hodges, Glenn Regehr, Mark Hanson, and Nancy McNaughton. Validation of an objective structured clinical examination in psychiatry. *Academic Medicine*, 73(8):910–912, 1998.

[HRJW97] R. L. Hulsman, W. J. G. Ros, M. Janssen, and A. M. Winnubst. INTERACT-CANCER. The development and evaluation of a computer-assisted course on communication skills for medical specialists in oncology. *Patient Education and Counseling*, 30(2):129–141, 1997.

[HRR96] D. Hong, G. Regehr, and R. K. Reznick. The efficacy of a computer-assisted preoperative tutorial for clinical clerks. *Canadian Journal of Surgery*, 39(3):221–224, 1996.

[HSH...86] K. J. Harvey, R. Stewart, M. Hemming, N. Naismith, and R. F. Moulds. Educational antibiotic advertising. *Medical Journal of Australia*, 145(1):28–32, 1986. Australia.

[HV97] H. Hoffman and D. Vu. Virtual reality: teaching tool of the twenty-first century? *Academic Medicine*, 72(12):1076–1081, 1997. United States.

[HVKA93] F. P. Hekelman, E. Vanek, K. Kelly, and S. Alemagno. Characteristics of family physicians' clinical teaching behaviors in the ambulatory setting: a descriptive study. *Teaching and Learning in Medicine*, 5:18–23, 1993.

[HW86] V. L. Hoellerich and Robert S. Wigton. Diagnosing pulmonary-embolism using clinical findings. *Archives of Internal Medicine*, 146(9):1699–1704, 1986.

[Irb92] D. M. Irby. How attending physicians make instructional decisions when conducting teaching rounds. *Academic Medicine*, 67(10, October):630–638, 1992.

[Irb94a] D. M. Irby. 3 exemplary models of case-based teaching. *Academic Medicine*, 69(12):947–953, 1994.

[Irb94b] D. M. Irby. What clinical teachers in medicine need to know. *Academic Medicine*, 69(5):333–342, 1994.

[Irb95] D. M. Irby. Teaching and learning in ambulatory care settings - a thematic review of the literature. *Academic Medicine*, 70(10):898–931, 1995.

[Irv97] D. Irvine. The performance of doctors. ii: Maintaining good practice, protecting patients from poor performance [see comments]. *British Medical Journal*, 314(7094):1613–1615, 1997. England.

[Jen94] P. Jenkins. An evaluation of day release courses for general practice trainees in south west thames region. *Education for General Practice*, 5(3 (supp.)):s341–s372, 1994.

[JHCS97] J. S. Jones, S. J. Hunt, S. A. Carlson, and J. P. Seamon. Assessing bedside cardiologic examination skills using "Harvey," a cardiology patient simulator. *Academic Emergency Medicine*, 4(10):980–985, 1997.

[JJM...94] Penny Jennett, Deborah Jones, Terrill Mast, Kathleen Egan, and Martyn Hotvedt. The characteristics of self-directed learning. In David A. Davis and Robert D. Fox, editors, *The physician as learner - linking research to practice*. American Medical Association, 1994.

[JLHM94] M. E. Johnston, K. B. Langton, R. B. Haynes, and A. Mathieu. Effects of computer-based clinical decision support systems on clinician performance and patient outcome. a critical appraisal of research [see comments]. *Annals of Internal Medicine*, 120:135–142, 1994.

[JSB...95] A. Jacques, A. Sindon, A. Bourque, G. Bordage, and J. J. Ferland. Structured oral interview. one way to identify family physicians' educational needs. *Canadian Family Physician*, 41:1346–1352, 1995.

[JSM...96] J. J. M. Jansen, A. J. J. A. Scherpbier, J. C. M. Metz, R. P. T. M. Grol, Cees P. M. van der Vleuten, and J. J. Rethans. Performance-based assessment in continuing medical-education for general-practitioners - construct-validity. *Medical Education*, 30(5):339–344, 1996.

[KD97] S. R. Kisely and S. P. Donnan. Competencies for part ii of the examination for membership of the faculty of public health medicine. *Journal of Public Health Medicine*, 19(1):11–17, 1997.

[KDH...98] E. H. Kevelighan, S. R. G. Duffy, R. M. Harris, A. J. Cole, K. Tait, and J. R. Hartley. An innovative special study module utilizing computer-based learning in obstetrics and gynaecology. *Medical Teacher*, 20(5):442–444, 1998.

[Kes95] I. G. Kestin. A statistical approach to measuring the competence of anaesthetic trainees at practical procedures [see comments]. *British Journal of Anaesthesia*, 75(6):805–809, 1995. Comment in: Br J Anaesth 1996 Oct;77(4):562.

[Kes96] I. G. Kestin. Cusum: a statistical method to evaluate competence in practical procedures [letter; comment]. *British Journal of Anaesthesia*, 77(4):562–562, 1996. Comment on: Br J Anaesth 1995 Dec;75(6):805-9.

[KH95] Peter Kinderman and Gary Humphris. Clinical communication skills teaching: The role of cognitive schemata. *Medical Education*, 29(6):436–442, 1995.

[Kin96] Sue Kinn. Information management and technology (IM&T) - training provision in postgraduate medical-education. *Medical Education*, 30(5):367–370, 1996.

[Kir96] R. M. Kirk. Teaching the craft of operative surgery. *Annals of the Royal College of Surgeons of England*, 78(1):25–28, 1996.

[KKD...89] J. M. Keay, G. N. Kenny, D. W. Daly, R. H. Adams, and K. C. Calman. Information technology in postgraduate medical-education. *Medical Education*, 23(5):472–475, 1989.

[KKP...96] A. W. Kushniruk, D. R. Kaufman, V. L. Patel, Y. Levesque, and P. Lottin. Assessment of a computerized patient record system: A cognitive approach to evaluating medical technology. *M.D. Computing*, 13(5):406–415, 1996.

[Kle80] Arthur K. Kleinman. *Patients and Healers in the Context of Culture : an exploration of the border land between anthropology, medicine, and psychiatry.* Number 3 in Comparative studies of health systems and medical care. University of California Press, Berkeley, 1980.

[Kle89] GaryA. Klein. Recognition-primed decisions. *Advances in Man-Machine Systems Research*, 5:47–92, 1989.

[KLZ...89] M. P. Knudson, F. H. Lawler, S. C. Zweig, C. A. Moreno, M. C. Hosokawa, and Jr. Blake, R. L. Analysis of resident and attending physician interactions in family medicine. *Journal of Family Practice*, 28(6):705–709, 1989.

[KM92] D. R. Kelly and T. S. Murray. An assessment of vocational training. *Medical Education*, 26(5):402–405, 1992.

[KN72] B. M. Korsch and V. F. Negrete. Doctor-patient communication. *Scientific American*, 227(2):66–74, 1972.

[Kop71] J. A. Kopta. An approach to the evaluation of operative skills. *Surgery.*, 70:297–303, 1971.

[Kov97] G. Kovacs. Procedural skills in medicine: linking theory to practice. *Journal of Emergency Medicine*, 15(3):387–391, 1997.

[KP98] Andre W. Kushniruk and Vimla L. Patel. Cognitive evaluation of decision making processes and assessment of information technology in medicine. *International Journal of Medical Informatics*, 51(2-3):83–90, 1998.

[KPM98] A. W. Kushniruk, V. L. Patel, and A. A. J. Marley. Small worlds and medical expertise: implications for medical cognition and knowledge engineering. *International Journal of Medical Informatics*, 49(3):255–271, 1998.

[Kra82] S. K. Krackov. *Influence of site on ambulatory care residency education in internal medicine.* Ph.D. dissertation, University of Rochester, Rochester, NY, 1982.

[L...86] J. Levenstein et al. The patient-centred clinical method. 1. a model for the doctor-patient interaction in family medicine. *Family Practice*, 3:75–79, 1986.

[LBB...90] H. Leclere, M. D. Beaulieu, G. Bordage, A. Sindon, and M. Couillard. Why are clinical problems difficult - general-practitioners opinions concerning 24 clinical problems. *Canadian Medical Association Journal*, 143(12):1305–1315, 1990.

[Les85] A. L. Lesser. Problem-based interviewing in general practice: a model. *Medical Education*, 19(4):299–304, 1985. England.

[LF84] III. Lippert, F. G. and J. A. Farmer. *Psychomotor skills in orthopedic surgery.* Williams & Wilkins, Baltimore, 1984.

[LH93] Huw Llewelyn and Anthony Hopkins. *Analysing how we reach clinical decisions.* Royal College of Physicians of London, London, 1993.

[LH94] Jocelyn Lockyer and Van Harrison. Performance assessment: the role of chart review analysis. In David A. Davis and Robert D. Fox, editors, *The physician as learner - linking research to practice.* American Medical Association, 1994.

[LHB88] Pawel Lewicki, Thomas Hill, and Elizabeth Bizot. Acquisition of procedural knowledge about a pattern of stimuli that cannot be articulated. *Cognitive Psychology*, 20:24–37, 1988.

[LHG...92] A. G. Lossing, E. M. Hatswell, T. Gilas, R. K. Reznick, and L. C. Smith. A technical-skills course for 1st-year residents in general surgery: a descriptive study. *Canadian Journal of Surgery*, 35(5):536–540, 1992.

[Lip93] R. Lipshitz. Converging themes in the study of decision making in realistic settings. In Gary A. Klein, Judith Orasanu, R. Calderwood, and C. E. Zsambok, editors, *Decision Making in Action : Models and Methods*, pages 103–137. Ablex, Norwood, NJ, 1993.

[LJR97] L. M. Lucas, S. K. Joos, and J. B. Reuler. A housestaff ambulatory block model - description and evaluation. *Teaching and Learning in Medicine*, 9(2):136–143, 1997.

[LK92] Alan Lesgold and Sandra Katz. Models of cognition and educational technologies: implications for medical training. In David A. Evans and Vimla L. Patel, editors, *Advanced Models of Cognition for Medical Training and Practice*, pages 255–265. Springer-Verlag, Berlin Heidelberg, 1992. Proceedings of the NATO Advanced Research Workshop on Advanced Models of Cognition for Medical Training and Practice held at Il Ciocco, Barga, Italy June 19-22 1991.

[LL98] Svein-Ivar Lillehaug and Susanne P. Lajoie. AI in medical education — another grand challenge for medical informatics. *Artificial Intelligence in Medicine*, 12:197–225, 1998.

[LM91] D. Lowdermilk and W. C. McGaghie. Resident learning in ambulatory care: skill adaptation and faculty supervision. *Qualitative Health Res.*, 1:100–116, 1991.

[Low86] D. L. Lowdermilk. *Environmental factors influencing ambulatory clinical learning experiences of internal medicine residents.* Ph.D. dissertation, University of North Carolina, Chapel Hill, 1986.

[LPC95] P. Lanyon, D. Pope, and P. Croft. Rheumatology education and management skills in general practice: a national study of trainees. *Annals of the Rheumatic Diseases*, 54(9):735–739, 1995.

[LRF...88] Alan Lesgold, Harriet Rubinson, Paul Feltovitch, Robert Glaser, Dale Klopfer, and Yen Wang. Expertise in a complex skill: Diagnosing x-ray pictures. In Michelene T. H. Chi, Robert Glaser, and M. J. Farr, editors, *The Nature of Expertise*, pages 311–342. Lawrence Erlbaum Associates, Publishers, Hillsdale, New Jersey, 1988.

[LT92] Richard J. Lilford and J. D. Thornton. Decision logic in medical practice, the Milroy Lecture 1992. *Journal of the Royal College of Physicians of London*, 26(4):400–412, 1992.

[LT97] J. Littlefield and C. Terrell. Improving the quality of resident performance appraisals. *Academic Medicine*, 72(10 Suppl 1):S45–S47, 1997. United States.

[Luc96] E. A. Luce. The plastic surgery teaching conference: a horse and buggy educational vehicle for supersonic times [editorial]. *Plastic & Reconstructive Surgery*, 97(6):1234–1237, 1996.

[LWLW96] M. Lloyd-Williams and F. Lloyd-Williams. Palliative care teaching and today's general practitioners — is it adequate? *European Journal of Cancer Care (English)*, 5(4):242–245, 1996.

[Maa90] Maatsch. Linking theories of competence to the construct validity of performance assessment tests. In Sherry L. Willis and Samuel S. Dubin, editors, *Maintaining Professional Competence : Approaches to Career Enhancement, Vitality, and Success Throughout a Work Life*, pages 95–124. Jossey-Bass Publishers, San Francisco, 1990.

[Mag92] Lorenzo Magnani. Abductive reasoning: philosophical and educational perspectives in medicine. In David A. Evans and Vimla L. Patel, editors, *Advanced Models of Cognition for Medical Training and Practice*, pages 21–42. Springer-Verlag, Berlin Heidelberg, 1992. Proceedings of the NATO Advanced Research Workshop on Advanced Models of Cognition for Medical Training and Practice held at Il Ciocco, Barga, Italy June 19-22 1991.

[Mar91] M. K. Marvel. Improving clinical teaching skills using the parallel process model. *Family Medicine*, 23(4):279–284, 1991.

[MB97] G. A. Mooney and J. G. Blight. Information technology in medical education: current and future applications. *Postgraduate Medical Journal*, 73:701–704, 1997.

[MBO91] M. Mugford, P. Banfield, and M. O'Hanlon. Effects of feedback of information on clinical practice: a review [see comments]. *British Medical Journal*, 303(6799):398–402, 1991. England.

[MC97] T. S. Murray and L. M. Campbell. Finance, not learning needs, makes general practitioners attend courses: a database survey [see comments]. *British Medical Journal*, 315(7104):353–353, 1997. England.

[MCR...97] H. M. MacRae, R. Cohen, G. Regehr, R. Reznick, and M. Burnstein. A new assessment tool: The patient assessment and management examination. *Surgery*, 122(2):335–343, 1997.

[McW85] Ian R. McWhinney. Patient-centred and doctor-centred models of clinical decision making. In Michael G. Sheldon, John Brook, and Alan Rector, editors, *Decision making in general practice*. Macmillan, London, 1985.

[MDZ...95] G. Mari, P. Denardi, A. Zerbi, G. Balzano, L. Zannini, A. Marassi, and V. Dicarlo. A postgraduate teaching course in laparoscopic surgery. *Surgical Endoscopy-Ultrasound and Interventional Techniques*, 9(10):1119–1122, 1995.

[MF88] P. Maguire and A. Faulkner. Improve the counselling skills of doctors and nurses in cancer care. *British Medical Journal*, 297(6652):847–849, 1988.

[MGAC96] A. A. Milne, J. M. T. Griffiths, R. J. Aitken, and T. J. Crofts. Comparative assessment of surgical training using objective criteria. *Annals of the Royal College of Surgeons of England*, 78(4 Supplement):177–179, 1996.

[MI97] S. R. McGee and D. M. Irby. Teaching in the outpatient clinic - practical tips. *Journal of General Internal Medicine*, 12(2):S34–S40, 1997.

[Miz84] T. Mizrahi. Managing medical mistakes: ideology, insularity and accountability among internists-in-training. *Social Science and Medicine*, 19:135–146, 1984.

[MJ93] M. L. Malone and T. C. Jackson. Educational characteristics of ambulatory morning report. *Journal of General Internal Medicine*, 8(9):512–514, 1993.

[MJWH97] H. Mandin, A. Jones, W. Woloschuk, and P. Harasym. Helping students learn to think like experts when solving clinical problems. *Academic Medicine*, 72(3):173–179, 1997.

[MKK88] Alan J. Moskowitz, Benjamin J. Kuipers, and Jerome P. Kassirer. Dealing with uncertainty, risks, and tradeoffs in clinical decisions - a cognitive science approach. *Annals of Internal Medicine*, 108(3):435–449, 1988.

[MLH98] W. S. McDougal, M. E. Lunz, and G. Hirst. Postgraduate education: does it improve the knowledge base of practitioners with time? *Journal of Urology*, 160(2):502–504, 1998.

[MM93] M. K. Marvel and P. K. Morphew. Levels of family involvement by resident and attending physicians. *Family Medicine*, 25(1):26–30, 1993.

[MMF...89] D. Mant, C. McKinlay, A. Fuller, T. Randall, E. M. Fullard, and J. Muir. Three year follow up of patients with raised blood pressure identified at health checks in general practice. *British Medical Journal*, 298(6684):1360–1362, 1989. England.

[MMT...93] R. Macura, K. Macura, V. Toro, E. Binet, and J. Trueblood. Case-based tutor for radiology. In H. U. Lemke, K. Inamura, C. C. Jaffe, and R. Felix, editors, *Computer Assisted Radiology, Proceedings of the International Symposium CAR'93*, pages 583–588. Springer-Verlag, 1993.

[MNA98] D. C. Macarthur, S. J. Nixon, and R. J. Aitken. Avoidable deaths still occur after large bowel surgery. Scottish audit of surgical mortality, Royal College of Surgeons of Edinburgh. *British Journal of Surgery*, 85(1):80–83, 1998.

[MPF98] I. R. Matthew, D. J. Pollard, and J. W. Frame. Development and evaluation of a computer-aided learning package for minor oral surgery teaching. *Medical Education*, 32:89–94, 1998.

[MPM...90] R. G. McAuley, W. M. Paul, G. H. Morrison, R. F. Beckett, and C. H. Goldsmith. Five-year results of the peer assessment program of the

College of Physicians and Surgeons of Ontario. *Canadian Medical Association Journal*, 143(11):1193–1199, 1990. Canada.

[MRA...94] S. Mahood, R. Rojas, D. Andres, C. Zagozeski, G. White, and T. Bradel. Educational contracts in family medicine residency training. *Canadian Family Physician*, 40:549–6, 1994. Canada.

[MRR...97] J. A. Martin, G. Regehr, R. Reznick, H. MacRae, J. Murnaghan, C. Hutchison, and M. Brown. Objective structured assessment of technical skill (OSATS) for surgical residents. *British Journal of Surgery*, 84(2):273–278, 1997.

[MTWS91] F. Morris, S. G. Tordoff, D. Wallis, and D. V. Skinner. Cardiopulmonary resuscitation skills of preregistration house officers: five years on [see comments]. *British Medical Journal*, 302(6777):626–627, 1991. England.

[MvKMP91] T. J. Meyer, D. van Kooten, S. Marsh, and A. V. Prochazka. Reduction of polypharmacy by feedback to clinicians. *Journal of General Internal Medicine*, 6(2):133–136, 1991. United States.

[MW96] J. G. Miller and F. M. Wolf. Strategies for integrating computer-based activities into your educational environment: A practical guide. *Journal of the American Medical Informatics Association*, 3(2):112–117, 1996.

[MWJS88] Marina Myles-Worsley, William J. Johnston, and Margaret A. Simons. The influence of expertise on x-ray image processing. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 14(3):553–557, 1988.

[MWM78] T. J. McGlynn, J. B. Wynn, and R. F. Munzenrider. Resident education in primary care: how residents learn. *Journal of Medical Education*, 53(12):973–981, 1978.

[MWMT92] T. R. Malloy, Robert S. Wigton, J. Meeske, and T. G. Tape. the influence of treatment descriptions on advance medical directive decisions. *Journal of the American Geriatrics Society*, 40(12):1255–1260, 1992.

[Nat99] National Health Service Executive. *Clinical governance in the new NHS.* Department of Health, London, 1999.

[NBC...85] Geoff Norman, Georges Bordage, Lynn Curry, Dale Dauphinee, Brian Jolly, David Newble, Arthur Rothman, Betsy Stalenhoef, Paula Stillman, David Swanson, and Xenia Tonesk. A review of recent innovations in assessment. In Richard Wakeford, Philip Bashook, Brian Jolly, and Arthur Rothman, editors, *Directions in clinical assessment.* 1st edition, 1985.

[NBC...96] G. R. Norman, L. R. Brooks, J. P. W. Cunnington, V. Shali, M. Marriott, and G. Regehr. Expert-novice differences in the use of history and visual information from patients. *Academic Medicine*, 71(10, S):S62–S64, 1996.

[NBR...96] G. R. Norman, L. R. Brooks, G. Regehr, M. Marriott, and V. Shali. Impact of feature interpretation on medical student diagnostic performance. *Academic Medicine*, 71(1, S):S108–S109, 1996.

[NCBB92] G. R. Norman, C. L. Coblentz, L. R. Brooks, and C. J. Babcook. Expertise in visual diagnosis - a review of the literature. *Academic Medicine*, 67(10, S):S78–S83, 1992.

[NDL...93] G. R. Norman, David A. Davis, S. Lamb, E. Hanna, P. Caulford, and T. Kaigas. Competency assessment of primary care physicians as part of a peer review program. *Journal of the American Medical Association*, 270(9):1046–1051, 1993.

[New94] David Newble. Guidelines for the development of effective and efficient procedures for the assessment of clinical competence. In David Newble, Brian Jolly, and Richard Wakeford, editors, *The certification and recertification of doctors: issues in the assessment of clinical competence*, pages 92–105. Cambridge University Press, Cambridge, England, 1994.

[NKLT96] Calvin F. Nodine, Harold L. Kundel, Sherri C. Lauver, and Lawrence C. Toto. Nature of expertise in searching mammograms for breast masses. *Academic Radiology*, 3:1000–1006, 1996.

[NMHH93] D. B. Nash, L. E. Markson, S. Howell, and E. A. Hildreth. Evaluating the competence of physicians in practice: from peer review to performance assessment. *Academic Medicine*, 68(2):S19–S22, 1993.

[NN85] Victor R. Neufeld and Geoffrey R. Norman. *Assessing clinical competence*. Number 7 in Springer series on medical education. Springer Publishing, New York, 1985.

[Nor92] North of England Study of Standards and Performance in General Practice. Medical audit in general practice. i: Effects on doctors' clinical behaviour for common childhood conditions. [see comments]. *British Medical Journal*, 304(6840):1480–1484, 1992. England.

[NS98] G. R. Norman and S. I. Shannon. Effectiveness of instruction in critical appraisal (evidence-based medicine) skills: a critical appraisal. *Canadian Medical Association Journal*, 158(2):177–181, 1998.

[NSK...84] R. J. Napodano, B. L. Schuster, S. K. Krackov, L. E. Young, P. F. Griner, and R. Pratt. Use of private offices in education of residents in internal medicine. *Archives of Internal Medicine*, 144(2):303–305, 1984.

[NTF...85] G. R. Norman, P. Tugwell, J. W. Feightner, Linda J. Muzzin, and L. L. Jacoby. Knowledge and clinical problem-solving. *Medical Education*, 19(5):344–356, 1985.

[NWMM95] D. Nicol, J. Ward, R. McMullin, and D. Moss. Urological training in Australasia: perceptions of recent fellows and current trainees. *Australia and New Zealand Journal of Surgery*, 65(4):278–283, 1995.

[OAHH89] P. Owen, L. Allery, K. Harding, and T. Hayes. General practitioners' continuing education within and outside their practice. *British Medical Journal*, 299:238–240, 1989.

[OLS...95] D. Ota, B. Loftin, T. Saito, R. Lea, and J. Keller. Virtual reality in surgical education. *Computers in Biology and Medicine*, 25(2):127–137, 1995.

[OMKR98] M. O'Connor, R. McGraw, L. Killen, and D. Reich. A computer-based self-directed training module for basic suturing. *Medical Teacher*, 20(3):203–206, 1998.

[OS97] Dermot C. O'Riordan and Nick Shaper. Other specialties can learn from level of supervision of surgical training. *British Medical Journal*, 315(15 nov):1307–1307, 1997.

[OSW93] L. M. Osborn, J. R. Sargent, and S. D. Williams. Effects of time-in-clinic, clinic setting, and faculty supervision on the continuity clinic experience. *Pediatrics*, 91(6):1089–1093, 1993.

[P...84] David Pendleton et al. *The consultation: an approach to learning and teaching.* Number 6 in Oxford general practice series. Oxford University Press, Oxford, 1984.

[PAAS98] V. L. Patel, V. G. Allen, J. F. Arocha, and E. H. Shortliffe. Representing clinical guidelines in GLIF: Individual and collaborative expertise. *Journal of the American Medical Informatics Association*, 5(5):467–483, 1998.

[Pat98] V. L. Patel. Individual to collaborative cognition: a paradigm shift? *Artificial Intelligence in Medicine*, 12(2):93–96, 1998.

[PB80] D. A. Pendleton and S. Bochner. The communication of medical information in general practice consultations as a function of patients' social class. *Soc. Sci. Med. [Med. Psychol. Med. Sociol].*, 14A:669–673, 1980.

[PB95] G. Page and Georges Bordage. The Medical Council of Canada's key features project: a more valid written examination of clinical decision-making skills [see comments]. *Academic Medicine*, 70(2):104–110, 1995. Comment in: Acad Med 1995 Feb;70(2):89-90.

[PBA95] G. Page, G. Bordage, and T. Allen. Developing key-feature problems and examinations to assess clinical decision-making skills. *Academic Medicine*, 70(3):194–201, 1995.

[PBD...95] G. G. Page, J. Bates, S. M. Dyer, D. R. Vincent, G. Bordage, A. Jacques, A. Sindon, T. Kaigas, G. R. Norman, M. Kopelow, and J. Moran. Physician-assessment and physician-enhancement programs in canada. *Canadian Medical Association Journal*, 153(12):1723–1728, 1995.

[PC89] R. W. Putnam and M. D. Campbell. Competence. In R. D. Fox, Paul Mazmanian, and R. W. Putnam, editors, *Change and learning in the lives of physicians.* Praeger Publishing, New York, NY, 1989.

[PC96] John Pitts and Colin Coles. Changes in teaching behaviour one year after a general practice trainers' course: implications for course planning. *Education for General Practice*, 7:199–207, 1996.

[PCS89] G. A. Paccione, E. Cohen, and C. E. Schwartz. From forms to focus. a new teaching model in ambulatory medicine. *Archives of Internal Medicine*, 149(11):2407–2411, 1989.

[PEK90] V. L. Patel, D. A. Evans, and D. R. Kaufman. Reasoning strategies and the use of biomedical knowledge by medical-students. *Medical Education*, 24(2):129–136, 1990.

[PG91] V. L. Patel and G. J. Groen. Developmental accounts of the transition from medical-student to doctor - some problems and suggestions. *Medical Education*, 25(6):527–535, 1991.

[PGAC96] M. A. Potter, J. M. Griffiths, R. J. Aitken, and T. J. Crofts. An objective assessment of surgical training. *Annals of the Royal College of Surgeons of England*, 78(1):11–13, 1996.

[PGF86] V. L. Patel, G. J. Groen, and C. H. Frederiksen. Differences between medical-students and doctors in memory for clinical cases. *Medical Education*, 20(1):3–9, 1986.

[PGN91] V. L. Patel, G. J. Groen, and G. R. Norman. Effects of conventional and problem-based medical curricula on problem-solving. *Academic Medicine*, 66(7):380–389, 1991.

[PGS88] V. L. Patel, G. J. Groen, and H. M. Scott. Biomedical knowledge in explanations of clinical problems by medical-students. *Medical Education*, 22(5):398–406, 1988.

[Pie93a] M. Pietroni. The assessment of competence in surgical trainees. *Annals of the Royal College of Surgeons of England*, 75(6, Suppl):200–202, 1993.

[Pie93b] M. Pietroni. Guidelines and standards in surgical training. *Annals of the Royal College of Surgeons of England*, 75(5):305–307, 1993.

[PKA95] V. L. Patel, D. R. Kaufman, and J. F. Arocha. Steering through the murky waters of a scientific conflict -situated and symbolic models of clinical cognition. *Artificial Intelligence in Medicine*, 7(5):413–438, 1995.

[PL99] E. Paice and P. Leaver. Improving the training of shos [editorial]. *British Medical Journal*, 318(7190):1022–1023, 1999.

[PMS...97] R. M. Poses, D. K. McClish, W. R. Smith, Mark Chaput de Saintonge, E. C. Huber, F. L. W. Clemo, B. P. Schmitt, D. AlexanderForti, E. M. Racht, C. C. Colenda, and R. M. Centor. Physicians' judgments of the risks of cardiac procedures -differences between cardiologists and other internists. *Medical Care*, 35(6):603–617, 1997.

[Pol96] A. V. Pollock. How do we measure surgical competence? [editorial]. *European Journal of Surgery*, 162(5):355–360, 1996.

[Pri98] M. Pringle. Preventing ischaemic heart disease in one general practice: from one patient, through clinical audit, needs assessment, and commissioning into quality improvement [see comments]. *British Medical Journal*, 317(7166):1120–discussion1124, 1998. England.

[PSA96] F. J. Papa, R. C. Stone, and D. G. Aldrich. Further evidence of the relationship between case typicality and diagnostic performance: Implications for medical education. *Academic Medicine*, 71(1, S):S10–S12, 1996.

[PSM90] Frank J. Papa, Jay H. Shores, and Steve Meyer. Effects of pattern matching, pattern discrimination, and experience in the development of diagnostic expertise. *Academic Medicine*, 65(9 Suppl, Supp):S21–S22, 1990. United States.

[PTODM94] H. M. Pieters, F. W. W. M. Touw-Otten, and R. A. De Melker. Simulated patients in assessing consultation skills of trainees in general practice vocational training: a validity study. *Medical Education*, 28(3):226–233, 1994.

[PWC...90] Nancy S. Palchik, Frederick M. Wolf, James T. Cassidy, Robert W. Ike, and Wayne K. Davis. Comparing information-gathering strategies of medical-students and physicians in diagnosing simulated medical cases. *Academic Medicine*, 65(2):107–113, 1990.

[PWC...97] E. Paice, G. West, R. Cooper, V. Orton, and A. Scotland. Senior house officer training: is it getting better? a questionnaire survey [see comments]. *British Medical Journal*, 314(7082):719–720, 1997. England.

[RAP...95] I. E. Rolfe, J. M. Andren, S. Pearson, M. J. Hensley, and J. J. Gordon. Clinical competence of interns. *Medical Education*, 29:22–230, 1995.

[RBF...91] M. F. Rhoton, A. Barnes, M. Flashburg, A. Ronai, and S. Springman. Influence of anesthesiology residents' noncognitive skills on the occurrence of critical incidents and the residents' overall clinical performances. *Academic Medicine*, 66(6):359–361, 1991.

[Ree93] M. W. R. Reed. Evaluation of surgical training - urgent improvement needed. *Annals of the Royal College of Surgeons of England*, 75:198s–199s, 1993.

[REM98] Eric Raufaste, Hlne Eyrolle, and Claudette Marin. Pertinence generation in radiological diagnosis: spreading activation and the nature of expertise. *Cognitive Science*, 22(4):517–546, 1998.

[RHK...95] D. L. Roter, J. A. Hall, D. E. Kern, L. R. Barker, K. A. Cole, and R. P. Roca. Improving physicians' interviewing skills and reducing patients' emotional distress. a randomized clinical trial. *Archives of Internal Medicine*, 155(17):1877–1884, 1995.

[RKS95] Lawrence Recht, Philip Kramer, and William J. Schwartz. Morning report in the computer era: tradition meets technology. *Medical Teacher*, 17(3):327–334, 1995.

[RL97] B. Reilly and M. Lemon. Evidence-based morning report: A popular new format in a large teaching hospital. *American Journal of Medicine*, 103(5):419–426, 1997.

[RN96] Glenn Regehr and Geoffrey R. Norman. Issues in cognitive psychology: Implications for professional education. *Academic Medicine*, 71(9):988–1001, 1996.

[Rog95a] E. Rogers. Via-rad: a blackboard-based system for diagnostic radiology. visual interaction assistant for radiology. *Artificial Intelligence in Medicine*, 7(4):343–360, 1995. Netherlands.

[Rog95b] Everett Mitchell Rogers. *Diffusion of Innovations*. Free Press, New York, 4th edition, 1995.

[Rog96] Erika Rogers. A study of visual reasoning in medical diagnosis. In Garrison W. Cottrell, editor, *Proceedings of the eighteenth annual conference of the cognitive science society*, pages 213–218. Lawrence Erlbaum Associates, Publishers, Mahwah, New Jersey, 1996.

[Roy92] Royal College of Radiologists Working Party. Influence of the Royal College of Radiologists' guidelines on hospital practice: a multicentre study. *British Medical Journal*, 304(6829):740–743, 1992. England.

[RRM...97] R. Reznick, G. Regehr, H. MacRae, J. Martin, and W. McCulloch. Testing technical skill via an innovative "bench station" examination. *American Journal of Surgery*, 173(3):226–230, 1997.

[RRYH98] D. A. Rogers, G. Regehr, K. A. Yeh, and T. R. Howdieshell. Computer assisted learning versus a lecture and feedback seminar for teaching a basic surgical technical skill. *American Journal of Surgery*, 175(6):508–510, 1998.

[RSD...91] Jan-Joost Rethans, Ferd Sturmans, Riet Drop, Cees P. M. van der Vleuten, and Pie Hobus. Does competence of general practitioners predict their performance? comparison between examination setting and actual practice. *British Medical Journal*, 303:1377–1380, 1991.

[RSFC97] A. M. Roche, R. W. Sanson-Fisher, and J. Cockburn. Training experiences immediately after medical school. *Medical Education*, 31(1):9–16, 1997.

[RT96] Elizabeth Rink and Jo Tissier. Long-term effect of minor surgery training workshops for general practice registrars. *Medical Education*, 30:466–467, 1996.

[RvLD...90] Jan-Joost Rethans, Yvonne van Leeuwen, Riet Drop, Cees P. M. van der Vleuten, and Ferd Sturmans. Performance and competence: two different constructs in the assessment of quality of medical care. *Family Practitioner*, 7:168–174, 1990.

[RWC89] P. Ramsden, G. Whelan, and D. Cooper. Some phenomena of medical students' diagnostic problem solving. *Medical Education*, 23:108–117, 1989.

[S...92] Eduardo Salas et al. Toward an understanding of team performance and training. In Robert W. Swezey and Eduardo Salas, editors, *Teams : Their Training and Performance*. Ablex Pub Corp, Norwood, N.J., 1992.

[S...93] Daniel Sulmasy et al. A randomized trial of ethics education for medical house officers. *Journal of Medical Ethics*, 19:157–163, 1993.

[Sal95] R. Salter. The US residency programme–lessons for pre-registration house officer education in the UK? *Postgraduate Medical Journal*, 71(835):273–277, 1995.

[Sal97] C. Salisbury. The Australian quality assurance and continuing education program as a model for the reaccreditation of general practitioners in the United Kingdom. *British Journal of General Practice*, 47(418):319–322, 1997. England.

[SB93] Henk G. Schmidt and Henny P. A. Boshuizen. On acquiring expertise in medicine. *Educational Psychology Review*, 5(3):205–221, 1993.

[SBM81] L. J. Sandlow, P. G. Bashook, and J. A. Maxwell. Medical care evaluation: an experience in continuing medical education. *J. Med. Educ.*, 56:580–586, 1981.

[SBO...98] C. M. Sox, H. R. Burstin, E. J. Orav, A. Conn, G. Setnik, D. W. Rucker, P. Dasse, and T. A. Brennan. The effect of supervision of residents on quality of care in five university-affiliated emergency departments. *Academic Medicine*, 73(7):776–782, 1998.

[SBS...91] M. Simpson, R. Buckman, M. Stewart, P. Maguire, M. Lipkin, D. Novack, and J. Till. Doctor-patient communication: the Toronto consensus statement. *British Medical Journal*, 303(6814):1385–1387, 1991.

[SD98] L. Southgate and D. Dauphinee. Maintaining standards in British and Canadian medicine: the developing role of the regulatory body. *British Medical Journal*, 316(7132):697–700, 1998.

[SdBT...95] Mike Sharples, Ben du Boulay, D. Teather, B. A. Teather, N. Jeffery, and G. H. du Boulay. The MR tutor: Computer-based training and professional practice. In *Proceedings of World Conference on Artificial Intelligence and Education, AIED'95*, pages 429–436. Washington, USA, 1995.

[SDJ...93] D. A. Sloan, M. B. Donnelly, S. B. Johnson, R. W. Schwartz, and W. E. Strodel. Use of an objective structured clinical examination (OSCE) to measure improvement in clinical competence during the surgical internship [see comments]. *Surgery.*, 114(2):343–350, 1993. Comment in: Surgery 1994 Nov;116(5):945-6; discussion 350-1.

[SDS...95] R. W. Schwartz, M. B. Donnelly, D. A. Sloan, S. B. Johnson, and W. E. Strodel. The relationship between faculty ward evaluations, osce, and absite as measures of surgical intern performance. *American Journal of Surgery*, 169(4):414–417, 1995.

[SDS...96] D. A. Sloan, M. B. Donnelly, R. W. Schwartz, J. L. Felts, A. V. Blue, and W. E. Strodel. The use of objective structured clinical examination (osce) for evaluation and instruction in graduate medical education. *Journal of Surg. Res.*, 63(1):225–230, 1996.

[SEKR97] Harold C. Sox, Jack Ende, Mark A. Kelley, and Paul G. Ramsey. *Graduate Education in Internal Medicine: A Resource Guide to Curriculum Development.* Federated Council for Internal Medicine, Inc, Philadelphia, 1997.

[SFM80] R. Sanson-Fisher and P. Maguire. Should skills in communicating with patients be taught in medical schools? *Lancet*, 2(8193):523–526, 1980. England.

[SGFL92] D. P. Sulmasy, G. Geller, R. Faden, and D. M. Levine. The quality of mercy. caring for patients with 'do not resuscitate' orders [see comments]. *Journal of the American Medical Association*, 267(5):682–686, 1992. United States.

[Sha90] J. Shapiro. Parallel process in the family medicine system: issues and challenges for resident training. *Family Medicine*, 22(4):312–319, 1990. United States.

[SJHC97] J. E. Smith, A. P. F. Jackson, J. Hurdley, and P. J. M. Clifton. Learning curves for fibreoptic nasotracheal intubation when using the endoscopic video camera. *Anaesthesia*, 52(2):101–106, 1997.

[SJT...97] Mike Sharples, N. Jeffery, D. Teather, B. A. Teather, and G. H. du Boulay. A socio-cognitive engineering approach to the development of a knowledge-based training system for neuroradiology. In *Artificial Intelligence in Education: Knowledge and Media in Learning Systems, Proceedings of AI-ED97*, pages 402–409. IOS Press, Amsterdam, 1997.

[SKD96] M. G. B. Sim, M. Kamien, and M. R. Diamond. From novice to proficient general practitioner: a critical incident study. *Australian Family Physician*, 25(2):S59–S64, 1996.

[SKF...97] Sanjiv Shah, Sheila Knowlden, Robert Fisher, Mark Harris, and Guan Yeo. General practice trainees in the 'borderline' zone of competence: the experience of remediation at a regional general practice training and education centre. *Education for General Practice*, 8:102–111, 1997.

[SKRM97] H. B. Slotnick, A. F. Kristjanson, R. R. Raszkowski, and R. Moravec. How doctors learn: mechanisms of action. Paper Presented at the Annual meeting of the American Educational Research Association, April 1997.

[SLM89] R. J. C. Steele, J. R. C. Logie, and A. Munro. Technical training in surgery: the trainee's view. *British Journal of Surgery*, 76:1291–1293, 1989.

[Slo99] H. B. Slotnick. How doctors learn in the medical workplace. Paper Presented at the Annual Meeting of the American Educational Research Association, April 1999 1999.

[SM86] S. W. Shankel and E. L. Mazzaferri. Teaching the resident in internal medicine: Present practices and suggestions for the future. *Journal of the American Medical Association*, 256(6):725–729, 1986.

[SM97] D. P. Sulmasy and E. S. Marx. Ethics education for medical house officers: Long term improvements in knowledge and confidence. *Journal of Medical Ethics*, 23(2):88–92, 1997.

[SMBH...96] Henk G. Schmidt, Maureen Machiels-Bongaerts, Helene Hermans, Theodorus J. tenCate, Ruud Venekamp, and Henny P. A. Boshuizen. The development of diagnostic competence: comparison of a problem-based, an integrated and a conventional medical curriculum. *Academic Medicine*, 71(6):658–664, 1996.

[SNB90] Henk G. Schmidt, Geoffrey R. Norman, and Henny P. A. Boshuizen. A cognitive perspective on medical expertise - theory and implications. *Academic Medicine*, 65(10):611–621, 1990.

[SNF95] S. Shannon, G. Norman, and J. Falconer. *Evaluation methods: a resource handbook*. McMaster University, 1995.

[SPC...98] E. H. Shortliffe, V. L. Patel, J. J. Cimino, G. O. Barnett, and R. A. Greenes. A study of collaboration among medical informatics research laboratories. *Artificial Intelligence in Medicine*, 12(2):97–123, 1998.

[SRPH90] P. L. Stillman, M. B. Regan, M. Philbin, and H. L. Haley. Results of a survey on the use of standardized patients to teach and evaluate clinical skills [see comments]. *Academic Medicine*, 65(5):288–292, 1990. Comment in: Acad Med 1991 May;66(5):271-2.

[SRRH96] David L. Sackett, Scott W. Richardson, William Rosenberg, and R. Brian Haynes. *Evidence-based medicine: how to practice and teach EBM*. Churchill Livingstone, New York, Edinburgh, 1996.

[SS92] Robert W. Swezey and Eduardo Salas. *Teams : Their Training and Performance*. Ablex Pub Corp, Norwood, N.J., 1992.

[SSR...91] Paula Stillman, David Swanson, Mary Beth Regan, Mary M. Philbin, Virginia Nelson, Thomas Ebert, Bryson Ley, Thomas Parrino, Jeanette Shorey, Alfred Stillman, Elaine Alpert, Joel Caslowitz, David Clive, James Florek, Milton Hamolsky, Charles Hatem, Janice Kizirian, Richard Kopelman, David Levenson, Gilbert Levinson, Jack McCue, Henry Pohl, Fred Schiffman, Joel Schwartz, Michael Thane, and Marshall Wolf. Assessment of clinical skills of residents utilizing standardized

patients. a follow-up study and recommendations for application [see comments]. *Annals of Internal Medicine*, 114(5):393–401, 1991. Comment in: Ann Intern Med 1991 Jul 15;115(2):158-9.

[ST91] J. Shapiro and Y. Talbot. Applying the concept of the reflective practitioner to understanding and teaching family medicine. *Family Medicine*, 23(6):450–456, 1991. United States.

[ST96] A. Singleton and A. Tylee. Continuing medical education in mental illness: a paradox for general practitioners [see comments]. *British Journal of General Practice*, 46(407):339–341, 1996. England.

[ST98] David Snadden and Mary L. Thomas. Portfolio learning in general practice vocational training - does it work. *Medical Education*, 32:401–406, 1998.

[Sta99] Standing Committee on Postgraduate Medical and Dental Education. *Doctors and dentists: the need for a process of review.* Standing committee on postgraduate medical and dental education, London, 1999.

[T...85] D. Tuckett et al. *Meetings Between Experts: an approach to sharing ideas in medical instituions.* Tavistock Publications, London, 1985.

[TARB97] J. M. Tracey, B. Arroll, D. E. Richmond, and P. M. Barham. The validity of general practitioners' self assessment of knowledge: cross sectional study [see comments]. *British Medical Journal*, 315(7120):1426–1428, 1997. England.

[TD85] P. Tugwell and C. Dok. Medical record review. In Victor R. Neufeld and Geoffrey R. Norman, editors, *Assessing clinical competence*, number 7 in Springer series on medical education, pages 142–182. Springer Publishing, New York, 1985.

[TKW92] T. G. Tape, J. Kripal, and Robert S. Wigton. Comparing methods of learning clinical-prediction from case simulations. *Medical Decision Making*, 12(3):213–221, 1992.

[Tou71] Allen Tough. *The adult's learning projects : a fresh approach to theory and practice in adult learning.* Number 1 in Research in Education series. Ontario Institute for Studies in Education, Toronto, Canada, 1971.

[TR96] J. Tissier and E. Rink. Evaluation of minor surgery courses for general practitioners. *Medical Education*, 30:333–338, 1996.

[TS98] Joseph K. H. Tan and Sameul Sheps. *Health Decision Support Systems.* Aspen, Gaithersburg, Md, 1998.

[TSO96] G. J. Toogood, C. F. Stableforth, and T. S. O'Brien. Surgical skills of pre-registration house surgeons. *Annals of the Royal College of Surgeons of England*, 78(3, Suppl):114–115, 1996.

[TW86] D. Tannen and C. Wallat. Medical professionals and parents - a linguistic analysis of communication across contexts. *Language in Society*, 15(3):295–311, 1986.

[TW87] D. Tannen and C. Wallat. Interactive frames and knowledge schemas in interaction -examples from a medical-examination interview. *Social Psychology Quarterly*, 50(2):205–216, 1987.

[UBS94] J. A. Ullian, C. J. Bland, and D. E. Simpson. An alternative approach to defining the role of the clinical teacher. *Academic Medicine*, 69(10):832–838, 1994.

[vdV96] Cees P. M. van der Vleuten. Making the best of the long case [see comments]. *Lancet*, 347(9003):704–705, 1996. England.

[vdVN94] Cees P. M. van der Vleuten and David Newble. Methods of assessment in certification. In David Newble, Brian Jolly, and Richard Wakeford, editors, *The certification and recertification of doctors: issues in the assessment of clinical competence*, pages 105–125. Cambridge University Press, Cambridge, England, 1994.

[vdVS90] Cees P. M. van der Vleuten and D. B. Swanson. Assessment of clinical skills with standardized patients: state of the art. *Teaching and Learning in Medicine*, 2:58–76, 1990.

[vdWSB98] M. W. van de Wiel, H. G. Schmidt, and H. P. Boshuizen. A failure to reproduce the intermediate effect in clinical case recall. *Academic Medicine*, 73:894–900, 1998.

[VF98] Malcolm Valentine and Fiona French. The introduction of personal learning plans for general practitioners in Grampian region, Orkney and Shetland. In R. Spencer-Jones, editor, *Self-directed learning for GPs*, pages 526–529. 1998. Supplement to: Education for General Practice; v9 no.4 (supp.); s497-s556.

[VGVRGW95] C. F. Von Gunten, J. H. Von Roenn, W. Gradishar, and S. Weitzman. A hospice, palliative medicine rotation for fellows training in hematology-oncology. *Journal of Cancer Educ*, 10(4):200–202, 1995.

[VHD79] J. E. Verby, P. Holden, and R. H. Davis. Peer review of consultations in primary care: the use of audiovisual recordings. *British Medical Journal*, 1(6179):1686–1688, 1979. England.

[vLMP...95] Y. D. van Leeuwen, S. S. L. Mol, M. C. Pollemans, M. J. Drop, R. Grol, and Cees P. M. van der Vleuten. Change in knowledge of general practitioners during their professional careers. *Family Practice*, 12(3):313–317, 1995.

[vRMP...95] A. M. van Rij, J. R. McDonald, R. A. Pettigrew, M. J. Putterill, C. K. Reddy, and J. J. Wright. Cusum as an aid to early assessment of the surgical trainee. *British Journal of Surgery*, 82(11):1500–1503, 1995.

[WA95] J. A. White and P. Anderson. Learning by internal-medicine residents - differences and similarities of perceptions by residents and faculty. *Journal of General Internal Medicine*, 10(3):126–132, 1995.

[WAC...96] C. D. Wyte, S. L. Adams, J. A. Cabel, K. Pearlman, P. R. Yarnold, M. Morkin, K. A. Hott, and J. J. Mathews. Prospective evaluation of emergency-medicine instruction for rotating first-postgraduate-year residents. *Academic Emergency Medicine*, 3(1):72–76, 1996.

[War86] Marx W. Wartofsky. Clinical judgement, expert programs, and cognitive style: a counter-essay in the logic of diagnosis. *Journal of Medicine and Philosophy*, 11:81–92, 1986.

[War95] S. A. Wartman. Morning report revisited: a new model reflecting medical practice of the 1990s. *Journal of General Internal Medicine*, 10(5):271–272, 1995.

[WBMN90] Robert S. Wigton, Linda L. Blank, H. Monsour, and Jo A. Nicolas. Procedural skills of practicing gastroenterologists. A national survey of 700 members of the American College of Physicians. *Annals of Internal Medicine*, 113(7):540–546, 1990. United States.

[WBNT89] Robert S. Wigton, L. L. Blank, J. A. Nicolas, and T. G. Tape. Procedural skills training in internal medicine residencies - a survey of program directors. *Annals of Internal Medicine*, 111(11):932–938, 1989.

[WFML91] Albert W. Wu, Susan Folkman, Stephen J. McPhee, and Bernard Lo. Do house officers learn from their mistakes. *Journal of the American Medical Association*, 265(16):2089–2094, 1991.

[WGSR88] Jr. Williamson, H. A., J. K. Glenn, D. C. Spencer, and J. C. Reid. The development of clinical independence: resident-attending physician interactions in an ambulatory setting. *Journal of Family Practice*, 26(1):60–64, 1988.

[WGW...89] J. W. Williamson, P. S. German, R. Weiss, E. A. Skinner, and F. Bowes III. Health science information management and continuing education of physicians. a survey of U.S. primary care practitioners and their opinion leaders. *Annals of Internal Medicine*, 110(2):151–160, 1989. United States.

[Wig81a] Robert S. Wigton. Effect of lectures and increased experience in gastroenterology on examination scores of internal medicine residents. *Gastroenterology*, 80(3):601–604, 1981.

[Wig81b] Robert S. Wigton. A method for selecting which procedural skills should be learned by internal medicine residents. *Journal of Medical Education*, 56(6):512–517, 1981.

[Wig92] Robert S. Wigton. Training internists in procedural skills. *Annals of Internal Medicine*, 116(12 pt 2, Pt2):1091–1093, 1992.

[Wil93] D. H. Wilson. Education and training of preregistration house officers: the consultants' viewpoint. *British Medical Journal*, 306(6871):194–196, 1993. England.

[Wil97] Janet A. Wilson. Unsupervised surgical training: questionnaire study. *British Medical Journal*, 314(7097):1803–1804, 1997.

[Woo90] D. Woods. Modeling and predicting human error. In J. Hochburg and B. Huey, editors, *Human Performance Models for Computer-Aided Engineering*, volume Card, S. Academic Press, Boston, elkind, j edition, 1990.

[WPS92] S. M. Williams, B. R. Parry, and M. M. Schlup. Quality control: an application of the cusum. *British Medical Journal*, 304(6838):1359–1361, 1992. England.

[WR93] S. Wheatly and A. D. Redmond. CPR and the RCP (1). training of doctors in NHS hospitals. *Journal of the Royal College of Physicians of London*, 27(4):408–411, 1993.

[WRCT94] C. P. Winckel, R. K. Reznick, R. Cohen, and B. Taylor. Reliability and construct validity of a structured technical skills assessment form. *American Journal of Surgery*, 167(4):423–427, 1994. United States.

[WS92] R. Wakeford and L. Southgate. Postgraduate medical education: modifying trainees study approaches by changing the examination. *Teaching and Learning in Medicine*, 4:210–213, 1992.

[WS93] N. S. Wenger and R. B. Shpiner. An analysis of morning report: implications for internal medicine education [see comments]. *Annals of Internal Medicine*, 119(5):395–399, 1993. Comment in: Ann Intern Med 1993 Sep 1;119(5):430-1; Comment in: Ann Intern Med 1994 Mar 1;120(5):442-3. United States.

[WSW95] R. Wakeford, L. Southgate, and V. Wass. Improving oral examinations: selecting, training, and monitoring examiners for the MRCGP. [see comments]. *British Medical Journal*, 311(7010):931–935, 1995. Comment in: BMJ 1995 Dec 9;311(7019):1572-3.

[WvdWG98] M. Wensing, T. van der Weijden, and R. Grol. Implementing guidelines and innovations in general practice: which interventions are effective? *British Journal of General Practice*, 48(427):991–997, 1998. England.

[XMD97] Y. Xiao, P. Milgram, and D. J. Doyle. Capturing and modeling planning expertise in anesthesiology: results of a field study. In Caroline E. Zsambok and Gary Klein, editors, *Naturalistic Decision Making (Expertise : Research and Applications)*, pages 197–206. Lawrence Erlbaum Associates, Mahwah, N.J., 1997.

[YHA...98] Steven R. Ytterberg, Ilene B. Harris, Sharon S. Allen, David C. Anderson, Philip M. Kofron, John H. Kvasnicka, James P. McCord, and James H. Moller. Clinical confidence and skills of medical students: use

of an osce to enhance confidence in clinical skills. *Academic Medicine*, 73(10 suppl):s103–s105, 1998.

[YLSH93] Michael J. Yedidia, Mack Lipkin, Mark D. Schwartz, and Colleen Hirschkorn. Doctors as workers: Work-hour regulations and interns' perceptions of responsibility. quality of care, and training. *Journal of General Internal Medicine*, 8:429–435, 1993.

[ZK97] Caroline E. Zsambok and Gary Klein. *Naturalistic Decision Making (Expertise : Research and Applications)*. Lawrence Erlbaum Associates, Mahwah, N.J., 1997.

Appendix A

Search Methodology

A search of PUBMED, BIDS and MEDLINE was conducted by **Ian Eiloart**, the researcher on the project.

PUBMED includes MEDLINE, PREMEDLINE, and other articles from Journals that provide full bibliographical information to Pubmed. around 3,900 current biomedical journals are indexed. A fuller description is available at
`http://www.ncbi.nlm.nih.gov/PubMed/overview.html`

"MeSH is a vocabulary of medical and scientific terms assigned to documents in PubMed by a team of experts. These terms can be used in many cases to search PubMed more efficiently than simple text words."

BIDS ISI indexes 7,500 journals in 3 databases. They do not say how many of them are indexed in the Science Citation Index.

Two on-line databases were searched for articles relevant to the study. PUBMED, which includes the MEDLINE database, `http://www.ncbi.nlm.nih.gov/PubMed/` indexes current biomedical journals, and was searched for articles published Sept 1988 - Sept 1998 with this search term:

```
        (MeSH term: "Education, Medical, Graduate")
    and (text word: "competence")
```

BIDS ISI Science Citation Index, a UK general science index, `http://www.bids.ac.uk/` was searched for articles published 1981 - Sept 1998 for these terms:

```
    (word in title, keywords or abstract: ("graduate" or "postgraduate"))
and (word in title, keywords or abstract: ("medical education")
```

In addition, the contents of Academic Medicine and the British Medical Journal were monitored as they came to publication throughout the period of the study.

The results of each search were then combined with our personal databases of references in professional (including medical) education and development and information technology.

Each search yielded in excess of 500 hits (many were duplicated), which were then

examined (by Ian Eiloart) on the basis of title and (where available) abstract for their relevance to the current project.

Additional articles were then found with searches on author names, where an author had published on the psychology of learning in Graduate Medical Education and via backwards citation searches.

Following this, Ian Eiloart and ourselves sorted the remaining articles into education topics and medical specialisms, which were labelled:

`aspects of judgement and competence; assessment; clinical experience; Information Technology; output measures; policy; supervision and feedback; teaching methods; theory`

`anaesthesia; dentistry; general practice; general medicine; obs & gyn; psychiatry; radiology; surgery`

Some articles were judged to be of peripheral interest, and are listed in our database although they did not form a part of the study. Others were judged to be irrelevant, for one of a variety of reasons, including that the material was:

- better described in another paper
- not relevant to current practice in the UK
- not chiefly about postgraduate medical experience
- not relevant to development of competence or judgement

This paper cites about 390 articles, leaving a further 470 uncited article citations in the database. Thus in all there are about 860 article citations in the wider database.

Appendix B

Consultations

The following senior doctors were consulted about attributes of competence and judgement, recent changes in postgraduate medical education and current issues of concern.

- Dr Justin Allen, Examiner for MRCGP and Secretary of the Joint Education Committee.
- Professor James Drife, (Leeds University), Junior Vice President Royal College of Obstetricians and Gynaecologists, Member of GMC Education Committee.
- Professor George du Boulay CBE, Emeritus Professor of Neurology at the Institute of Neurology.
- Dr Dennis Eraut, Chair of HST panel for Respiratory Medicine, North East Thames Region.
- Dr Noushin Farhoumand, Clinical Tutor in Psychiatry, South Downs Health NHS Trust.
- Dr Richard Gray, Associate Dean for General Practice, South Thames Region.
- Dr Neville Harrison, Associate Dean (surgery) South Thames Region.
- Dr Roger Neighbour, convenor for MRCGP examination.
- Dr Andrew Polmear, Trafford Centre for Medical Research, University of Sussex (concerned with Evidence-Based Medicine in Primary Care).
- Professor Lesley Southgate (CHIME, University College London), Leader of group who developed and are now implementing GMC Performance Procedures.
- Professor Richard Vincent, Director of Trafford Centre for Medical Research, University of Sussex and Consultant Cardiologist, Brighton Health Care.
- Professor Watson, (UMDS, London), Royal College of Psychiatry.

In addition, leading medical education researchers from Canada, U.S. and the Netherlands were consulted about research into Postgraduate Medical Education at conferences

in April 1998 and April 1999 in the U.S. and Canada (not at DoH expense); and their publications and citations were followed up. Also in April 1998, a 2 day visit was made to North Dakota Medical School to study their residency programmes and gain sufficient background knowledge to properly interpret North American research literature on the education of residents.