The MR Tutor: Computer-based Training and Professional Practice

M. Sharples, J.B.H. du Boulay
School of Cognitive and Computing Sciences, University of Sussex,
Brighton, BN1 9QH, U.K.
Tel: +44 (0) 1273 606755 Email: rad@cogs.susx.ac.uk

B.A. Teather, D. Teather, N. Jeffery
Medical Systems Research Group, School of Computing Sciences,
De Montfort University, Leicester, LE1 9BH, U.K.

G.H. du Boulay
(*Visiting Professor at the Medical Systems Research Group, De Montfort University*)
Institute of Neurology, Queen Square,
London, WC1N 3BG, U.K.

Abstract

The MR Tutor is intended for use by radiologists who are being trained to interpret MR images of the brain, particularly images presenting diseases that are acknowledged to be hard to differentiate. We have developed an image description language for MR images of the head suitable for a wide range of image sequences. The language is image-based, not pathology-based, and is used to describe the images as they present themselves rather than to classify the pathology which may be indicated in the image. We have established a dedicated archive of cases that illustrate a range of pathologies. This archive provides a substantial part of the domain knowledge for the MR Tutor. This paper is concerned with accounts of professional practice and skill development, and how these provide requirements for the design of the MR Tutor and similar systems.

Developing an effective computer-based system for professional training poses many additional problems compared to computer-aided instruction in an educational setting. A central difficulty is that of gaining the confidence of the trainee. The trainee must be convinced that it is worthwhile investing time in learning and operating the system, and the quality of training must be worth the effort. It is not enough just to offer useful training; it must also be provided in a way that fits the working life of a professional.

The MR Tutor project is addressing these issues through the development of a tutoring system for Magnetic Resonance (MR) imaging of the head. The system is designed to be used in a real medical setting and to be both endorsed by expert radiologists and useful to trainees. The project has brought together expertise in cognitive modelling, ITS design, radiology, medical imaging, and medical statistics. The aims are to:

- supplement professional training in radiology by offering computer-based training and diagnostic assistance;

- provide trainee radiologists with exposure to a large structured library of medical images and associated descriptions;
- help radiologists to develop a 'conceptual framework' and appropriate image description language for image interpretation.

We have already developed a detailed MR image description language suitable for a wide variety of image sequences. This language, with each case described in terms of a vector indicating the presence of absence of image features, provides the domain knowledge for the MR Tutor. The Tutor under-development deals with two distinct training issues: (i) training radiologists to view and describe an MR image set of the brain in a systematic manner, and (ii) training radiologists to make reliable diagnoses, especially in the case of confusable diseases.

The MR Tutor is currently being implemented using HiPWorks, a multimedia extension to the Poplog programming environment, running under X-Windows on Sun workstations.

This paper is concerned with the process of designing the MR Tutor, in particular with the ways in which accounts of professional practice and skill development are providing requirements for system design. A companion paper describes how theories of categorisation and concept teaching have provided us with a further set of guidelines for the design of the MR Tutor, and how these guidelines have been realised in the prototype system (Sharples et al., 1994).

1 Professionalism

Professionalism is both an ideology and a practice (Eraut, 1994). Seen as an ideology, a profession is a means of providing services that the general public are not competent to practice or evaluate. The emphasis is on careful recruitment to the profession, shared knowledge, and agreed standards of skill and behaviour. From the classroom to the police interview room, professionals must not only be competent, but also publicly accountable, and they need to protect themselves against misinformation and carelessness.

One means of providing accountability is to develop a more structured approach to the task, and to ensure that activities are recorded and can be evaluated by colleagues. Such an approach should lead to the establishment of guidelines for best practice based on a core of shared procedures and common terminology and thus restrict maverick behaviour.

As a practice, professionalism is situated in the confusion and uncertainty of the everyday world. Professionals sometimes work to tight schedules and cannot always afford to gather 'crystalline' evidence to support their decisions: "One whose work requires practical application to concrete cases simply cannot maintain the same frame of mind as the scholar or scientist." (Friedson, 1971 cited in Eraut, 1994).

This is the central tension of professionalism: the work is situated in a complex environment, relying on partial knowledge and requiring time-pressured decisions, but the results of that work must be made accountable in a well-structured form to clients, colleagues and the law. Nowhere is this tension stronger than in the medical professions.

One way to address this problem is to provide a shared 'conceptual framework' for the profession: a set of publicly available terms, structures and techniques that provide a means of communicating knowledge to colleagues, and a common language with which to record activities and justify decisions. The medical profession is devoting great effort to developing standard terminologies to describe symptoms, pathologies, and the content of medical images (e.g. Rector et al., 1994; du Boulay et al., 1994). The computer can play an important part by providing a means of storing and accessing the shared language, with the terms cross-referenced and accompanied by definitive examples.

Another approach to squaring the demands of professional practice with the need for competence and accountability is to promote the "reflective practitioner" (Schön, 1983), by encouraging professionals to reflect on their work and to develop strategies for evaluating and restructuring their actions. This can take place on the job (reflection-in-action), usually prompted by a breakdown in routine or an abnormal event, or after the job has finished when the professional attempts to make sense of the experience and learn from it. Again, the computer can contribute, by providing cases and events for comparison — along with strategies for repairing breakdowns — and an environment that simulates working practice but in a more structured, controllable form, to assist reflection and experiment.

Computer-based 'coached practice' environments such as SOPHIE (Brown et al., 1982), SHERLOCK (Lesgold et al., 1992), the Recovery Boiler Tutor (Woolf et al., 1986) and IDM (Fink, 1991) are examples of successful systems for training in complex skills. The theory of cognitive apprenticeship (Collins et al., 1989) that underlies these systems is appropriate to training in the professions, but it needs to be supplemented by a more specific set of requirements for system design. Eraut (1994) suggests that any framework for promoting and facilitating professional learning will have to take account of:

- an appropriate combination of learning settings (on-the-job, near the job, home, library, course, etc.);
- time for study, consultation and reflection;
- availability of suitable learning resources;
- people who are prepared (i.e. both willing and able) to give appropriate support;
- the learner's own capacity to learn and to take advantage of the opportunities available.

To this list we have added five more factors:

- the trainee's existing expertise;
- authentic and up-to-date material for study;
- appropriate presentation of learning materials (attractive, motivating, and neither patronising nor incomprehensible);
- the need to relate practice to learning, and to put the learning quickly into practice;
- the need to interleave learning with other activities that are competing for time and resources (e.g. professional work, administration, family);

So far, we have concentrated on the common aspects of professional work, but a successful training system must also be sensitive to the differences, between professions, levels of expertise, and individuals.

2 The Development of Professional Expertise

Two important areas of difference between professions are the application of knowledge and the type of practice. Radiologists develop organised knowledge schemas that enable rapid detection of abnormalities in an image. Lesgold et al. (1989) have shown that expert radiologists can make accurate interpretations after viewing an X-ray image for as little as two seconds.

But a radiologist must also be able to call on biomedical knowledge when discussing a case with colleagues, or when justifying an interpretation. In a series of experiments on the role of biomedical knowledge Boshuizen and Schmidt (1992) have shown that experts acquire a robust knowledge base that integrates general and situated knowledge. Knowledge integration is an active process requiring: (i) articulating a global framework (the biomedical knowledge), (ii) reflecting on situated experiences (individual cases as they are encountered), and (iii) actively making connections between situated knowledge and the global framework (Boshuizen and Schmidt, 1992). The integration of theoretical and experiential knowledge may be different in other professions, depending, for example, on the stability of the general knowledge and the time available for reflection.

Another significant difference between professions lies in the type of practice. Professional practices include diagnosis (e.g. physicians, electronic troubleshooters), design (e.g. architects, interior designers) and interpretive communication (e.g. teachers, actors, barristers). In the absence of a general theory of professional practice, this paper will concentrate on the task of radiologists, which is to describe the abnormal features of radiological images and provide an interpretation of them, in terms of underlying pathologies and anatomical structures.

Training should also be responsive to differing levels of expertise. Again, there are many differences between professions, but also some commonalities. A novice in any profession begins by acquiring background knowledge. In the case of radiology, this includes anatomy and general medicine. On starting to practice the profession, the trainee encounters many individual cases, generally in a work setting away from the textbook. A senior staff radiologist (in the U.K.) sees in the order of 80-100 cases per week, or around 100,000 in a working life. Through the experience of casework, a professional gains an extensive stock of mental schemas and performance knowledge. An expert is able to invoke rapidly a schema in the right problem area and can test this initial schema against the evidence (Lesgold et al., 1989). Lesgold suggests that, in the process of becoming an expert, a person acquires fragments of automatized procedure that gradually become integrated into extended sequences that guide performance. These sequences can be formed through the composition of fragments of activity, but they can also be taught explicitly, as a list of steps towards solution. An important issue in the training of professional skills is how to choose and teach routines and sequences of actions. Some training systems, in music and ballet for example, rely on well-developed sets of exercises and routines. This type of training is far less common in the medical profession.

The development of skill in radiology has the additional feature that there is a shift from visual to structural interpretation. Experts can imagine the 3-D structures that give rise to the 2-D radiographic images. They can identify directly features corresponding to anatomical parts, and become less sensitive to deviations, such as blemishes that are not diagnostically significant (Myles-Worsley, Johnston and Simons, 1988).

These findings have implications for the design of training systems for professional skills, and in particular for radiology. They suggest that a tutoring system should:

- base the training on cases derived from work practice;
- order the cases in a sequence that promotes understanding and retention;
- help the trainee to reflect on experience and to integrate general and situated knowledge (for example by relating cases to general principles or reference material);
- assist the trainee to integrate fragmentary knowledge into more general, guiding schemas;
- provide support for developing expertise, with a fading of support and a movement from a teaching to a work support system;
- provide ways for the trainee to store, annotate and index cases for future reference.

In radiology:

- assist the trainee to acquire structural schemas and to map images onto anatomical and pathological objects;
- help the trainee to make rapid, accurate judgements.

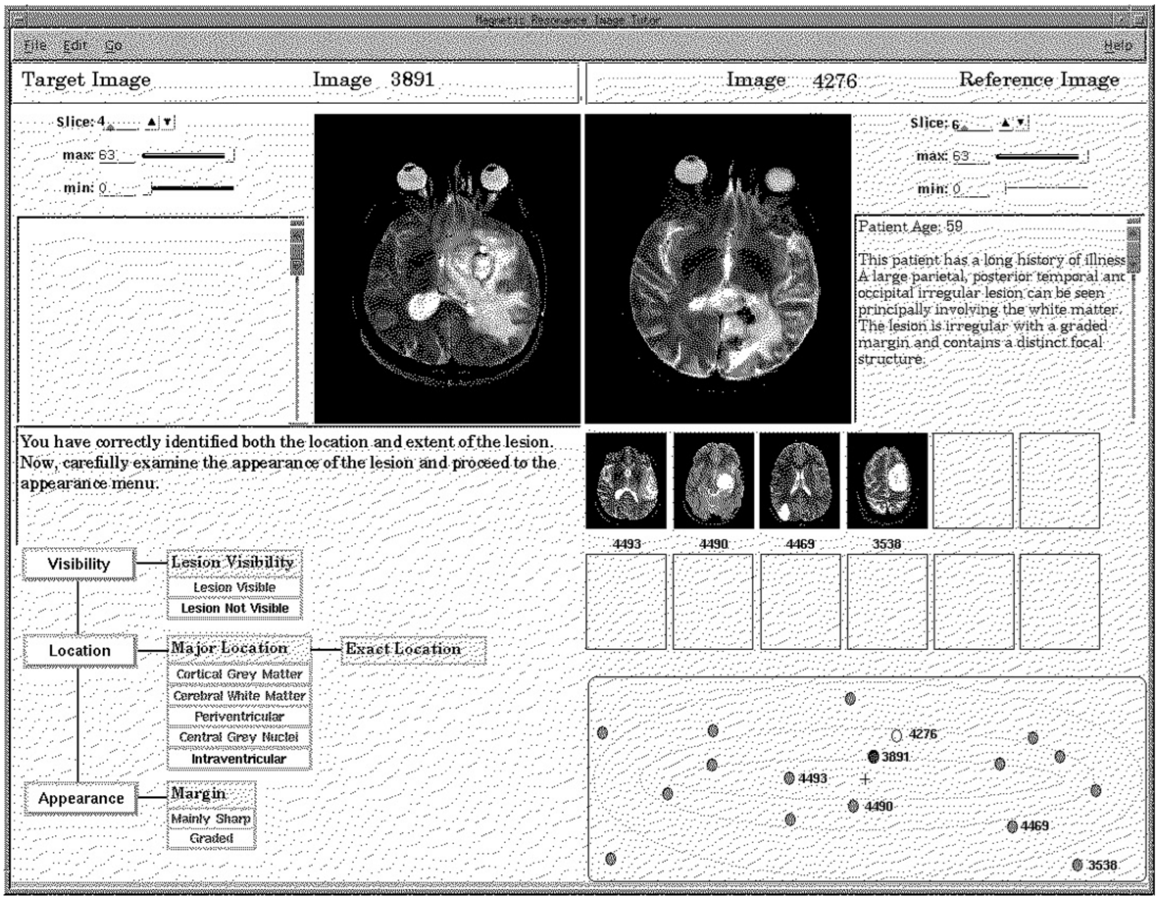
This is a demanding set of requirements for a computer-based tutor, and no existing system comes near to meeting them. The MR Tutor is an attempt to provide training in the interpretation of MR images that will be accepted as part of professional education. In the remainder of this paper we describe some aspects of the system, in terms of three of the guiding principles: the need for authentic and up to date material, appropriate presentation of learning materials, and support for developing expertise.

3 General design of the MR Tutor

The MR Tutor is intended for use by radiologists who are being trained to interpret MR images of the brain, particularly images presenting features that may be indicative of more than one disease. The training is in two stages. The first stage helps the trainee to acquire a structured language with which to describe the features of MR images. The second stage, not yet implemented, will help the trainee to discriminate between confusable cases from competing pathologies.

In the first stage of the tutor (see Figure 1), the trainee can browse through an archive of stacks of images for a given pathology and either the computer or the trainee can select a case for tutoring. The literature on categorization and concept teaching offers guidelines on how to order the presentation of exemplars to promote learning (see Sharples et al., 1994 for a discussion). The trainee then compares the target case with reference cases selected from the image archive. These reference images are presented with full descriptions, and the trainee works by making visual comparisons between target and reference images, and selecting terms to describe the target case taken from the structured description language.

Figure 1: Interface for image description



3.1 Authentic and up-to-date material

We have developed an *image description* language for MR images of the head suitable for a wide range of image sequences. The language is image-based, not pathology-based, and is used to describe the images as they present themselves rather than to classify the pathology which may be indicated in the image. Each stack of images is accompanied by a description and follow-up/biopsy/autopsy information to provide confirmed diagnoses. Statistical methods are then used to relate descriptions of images to confirmed diagnoses. The language has been incorporated into a computer tool (separate from the MR Tutor) that allows an expert rapidly to generate descriptions of sample images. Using this tool, the team has established a dedicated archive of some 1121 cases that illustrate a range of pathologies (Teather et al., 1994). This archive provides a substantial portion of the domain knowledge for the MR Tutor.

The terms of the language have been refined through a series of validating experiments, so as to provide the set of image feature descriptors that are necessary and sufficient to cover each pathology of interest and to distinguish it from competing pathologies. As the language was developed for an expert radiologist to create an annotated image archive, it has been simplified for tutoring purposes. Figure 2 shows a portion of the simplified language, associated with the shape and interior pattern of a lesion. An independent evaluation of an earlier system for CT imaging (Teather et al., 1988) has identified the training benefits of

```
Margin
     Mainly sharp
     Graded
Shape
     Rounded
     Irregular
             Area: sq. cm.
     Conforming to an anatomical feature
Interior Pattern
     Homogeneous
     Heterogeneous
     Containing a distinct focal structure
```

Figure 2: A portion of the Simplified Image Description Language relating to the appearance of a lesion

a structured approach to reporting, linked to a reference set of annotated example images and diagrams.

The image description language is particularly appropriate to computer-based tutoring, since it provides a canonical set of feature descriptors. These are being implemented in the MR Tutor as a frames system in which the features are represented by slots whose fillers indicate feature values.

3.2 Appropriate presentation of learning materials

There is an issue in providing tutoring material based on complex images, of how to give the trainee an overview of the entire image archive. Can we provide an index to the images based on visual appearance? How can the Tutor select images for teaching in an ordered sequence according to typicality, similarity and probability of category membership? This is part of a more general problem of storing and accessing visual material for picture archives, on-line reference libraries and multimedia presentations.

MR images for a particular 'sequence' (i.e. images generated according to a particular menu of gradient currents etc.) are organised into a stack where each image corresponds to a 'slice' through the head. A complete stack in a particular orientation will normally consist of more than twenty slices. A case will often consist of more than one stack of images generated under a different sequence or at a different orientation.

Each image stack in the MR Tutor archive is associated with a full expert description using the terms of the image description language. The set of descriptions form a 'feature space' where points in the space indicate exemplar cases and regions represent categories (pathologies). The statistical technique of Multiple Correspondence Analysis can take a high-dimensional space of nominal descriptors and display them as a two-dimensional scatter plot, or 'overview space', that represents as much of the variability of the feature space as possible. The technique makes no assumptions about independence of feature values. The space can be scaled so that the typicality of a case is proportional to its distance from the centre of the plot, and the similarity of any two cases is proportional to their nearness in the plot (Teather et al., 1994). By 'typicality' we mean 'central tendency' of a category rather than 'frequency of occurrence' of a case.

An example of an 'overview space' is shown at the lower right of Figure 1. This shows 17 cases for a particular pathology, where each case is indicated by a small circular point. The circle for the fairly typical case, currently the subject of tutoring (number 3891), is shown darker than the rest. The circle for a similar reference case (number 4276) is shown lighter than the rest. One image from the stack or series available for this case is displayed in the reference image window and was selected for display by 'clicking' on the point in the overview space.

The trainee can see at a glance which are typical cases (nearer the centre) and which are atypical or can select similar cases by directly clicking on adjacent points in the overview space for that pathology. The contours in the overview space indicate the degree of typicality. The training can thus incorporate appropriate exposure to both typical and atypical cases — the former otherwise sometimes being overshadowed by the latter in training (Schmidt and Boshuizen, 1993). When the trainee is choosing a feature value for an image (see lower left hand side of Figure 1) the overview space can light up those points which have the same feature value, thus providing a ready indication of which other images in the archive might be helpful. This facility can be organised for each descriptor independently, so the attempt to discriminate between descriptors can be supported by reference to all relevant images in the overview. The facility can also be used cumulatively to 'narrow down' the set of cases in the overview that conform to as much of the description as has so far been entered. Once the trainee has selected a complete set of descriptors, the Tutor can compute and display a point in the overview space to show how far in the space the trainee's description lies from that of the target case. The combination of a direct manipulation overview and menu selection of descriptors assists the trainee in making rapid accurate judgements. We believe that Multiple Correspondence Analysis and the overview space is a powerful general mechanism for interacting with picture archives.

3.3 Support for developing expertise

The MR Tutor supports the integration of general and situated knowledge by employing a structured image description language. This is not just a way of describing images via the computer, it also provides a grounded set of terms for the trainee to communicate with colleagues. Furthermore, by reflecting on each case in the terms of the description language, the trainee is building mental connections between situated experience and the stable knowledge structure provided by the language.

The tutor is also designed to match the trainee's developing expertise. We described earlier how the archive was built by an expert using a computer-based tool to annotate all the images of a large number of patients' brain lesions. This expert's tool is largely menu-driven and has extensive help facilities but does not provide all the extra scaffolding of the MR Tutor. For example, the tool does not provide an overview space and reference images are not readily accessible.

As the trainee develops in expertise, the tutor will move from the fully scaffolded training version shown in Figure 1 toward a version of the expert's tool. The scaffolding can be removed in stages of differing granularity, from switching off the lighting up of the points in the overview space when a description is being selected, to wholesale removal of the overview space.

In addition to scaffolding, the MR Tutor will provide tutorial feedback on the accuracy of the trainee's description and on the *process* of forming that description. Given that each image is fully described, tutoring can be directed, if need be, toward cases that present problematic features either because they are intrinsically difficult to describe accurately, or because they are difficult for a particular trainee. Under the heading of 'process' we include appropriate use of the image viewing controls, sufficient exploration of all the slices in the image stack before committing to a description, appropriate use of the archive and so on.

4 Conclusions

This paper has covered one aspect of the software design process of the MR Tutor. It has proposed a set of general principles, derived from theories of professional training and development of expertise, that can inform the design of training systems for professional practice. It has shown how three of these principles have informed the design of our system. The current system meets these three requirements, and we intend to incorporate the other principles into future versions of the Tutor. For example, we intend to:

- provide the trainee with facilities for study, consultation and reference through on-line access to case notes and reference materials;
- offer support for reflection on experience via personal annotations, notebooks and by annotating the overspace to form a visible 'student model';
- investigate the use of computer conferencing to provide expert support;
- integrate the Tutor with a computer-based brain atlas, to assist the trainee in mapping images onto anatomical objects.
- interleave training with work, by providing the Tutor as a window on the reporting consoles already in use for MR image capture and manipulation.

There is much work to be done in developing both the system and the principles as well as evaluating them in a variety of settings such as radiology departments of hospitals and medical libraries. The current implementation has provided enough functionality to conduct some initial evaluation with radiologists. The results of these initial studies are that the system can be used with minimal training, that the radiologists find the overview space to be a useful means of visualising and comparing the archive of cases, and that there is a clear need for a structured approach to describing MR images.

References

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

Brown, J.S., Burton, R.R., and de Kleer, J. (1982) Pedagogical, natural language and knowledge engineering techniques in SOPHIE I, II and III. In Sleeman, D. and J.S. Brown (eds.) *Intelligent Tutoring Systems*, pp. 227–282. London: Academic Press.

Collins, A., Brown, J.S. and Duguid, P. (1989) Cognitive apprenticeship: teaching the craft of reading, writing, and mathematics. In L. Resnick (ed.), *Knowing, Learning and Instruction: Essays in Honor of Robert Glaser*, pp. 453–494. Hillsdale, New Jersey: Lawrence Erlbaum.

du Boulay, G.H., Teather, B.A., Teather, D., Higgott, M.A. and Jeffery, N.P. (1994) Standard terminology for MR image description. *Proceedings of XV Symposium Neuroradiologium*, Kumamoto, Japan.

Eraut, M. (1994) *Developing Professional Knowledge and Competence*. London: The Falmer Press.

Fink, P. and Lusth. J. (1987) Expert systems and diagnostic expertise in the mechanical and electrical domains. *IEEE Transaction on Systems, Man, and Cybernetics*, **17** (3), 340–349.

Friedson, E. (1971) *Profession of of Medicine: A Study of Sociology of Applied Knowledge*. New York: Dodd, Mead and Co.

Lesgold, A., Glaser, R., Rubinson, H., Klopfer, D., Feltovich, P. and Wang, Y. (1989) Expertise in a complex skill: diagnosing X-ray pictures. In M. Chi, R. Glaser and M. Farr (eds.) *The Nature of Expertise*. Hillsdale, New Jersey: Lawrence Erlbaum.

Lesgold, A., Lajoie, S., Bunzo, M. and Eggan, G. (1992) SHERLOCK: A coached practice environment for an electronics troubleshooting job. In J. H. Larkin and R.W. Chabay (eds.) *Computer Assisted Instruction and Intelligent Tutoring Systems: Shared Goals and Complementary Approaches*, pp. 201–238. Hillsdale, New Jersey: Lawrence Erlbaum.

Myles-Worsley, M., Johnston, W.A. and Simons, M.A. (1988) The influence of expertise on X-ray processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **14** (3), 553–557.

Rector, A.L., Gangemi, A., Galeazzi, E., Glowinski, A.J. and Rossi-Mori, A. (1994) The Galen Model Schemata for anatomy: towards a reusable application-independent model of medical concepts. In P. Barahona, M. Veloso and J. Bryant (eds.) *Proceedings of the Twelfth International Congress of the European Federation for Medical Informatics*, MIE 94, Lisbon, pp. 229–233.

Schmidt, H.G. and Boshuizen, H.P.A. (1993) On acquiring expertise in medicine, *Educational Psychology Review*, **5** (3), 205–221.

Schön, D. (1983) *The Reflective Practitioner: Towards a New Design for Teaching and Learning in the Professions*. San Francisco: Jossey-Bass.

Sharples, M., du Boulay, B., Teather, D., Teather, B.A., Jeffery, N. and du Boulay, G.H. (1994) The Cognitive Basis for an MR Image Tutor. *Proceedings of East-West Conference on Computer Technologies in Education*, Part 1, pp. 214–219, Crimea.

Teather, D., Teather, B.A., Sharples, M., Jeffery, N., du Boulay, B., and du Boulay, G.H. (1994) Intelligent Tutoring for MR Imaging of the Head and Cerebral Disease. *Proceedings of Twelfth International Congress of the European Federation for Medical Informatics*, MIE 94, Lisbon.

Teather, D., Teather, B.A., Wills, K.M., du Boulay, G.H., Plummer, D., Isherwood, I. and Gholkar, A. (1988) Evaluation of Computer Advisor in the Interpretation of CT images of the Head. *Neuroradiology*, **30**, 511–517.

Woolf, B.P., Blegen, D., Jansen, J.H. and Verloop, A. (1986) Teaching a complex industrial process. *Proceedings of the National Conference on Artificial Intelligence*, Philadelphia, vol. II, pp. 722–728.