Commentary

Perception research in medical imaging

[1]D J MANNING, PhD, FInstP, [2]A GALE, PhD and [3]E A KRUPINSKI, PhD

[1]*Division of Medical Imaging Sciences, St. Martin's College Lancaster,* [2]*Institute of Behavioural Sciences, University of Derby, UK and* [3]*Department of Radiology Research, University of Arizona, Tuscon, AZ, USA*

The function of medical imaging is to provide information and to reduce diagnostic uncertainty. Radiological information takes many forms including the spatial distributions of attenuation coefficients, acoustic impedances, proton densities, radiopharmaceutical concentrations and so on. The precision with which the information is presented also continues to be refined so the images produced now give outstanding reproduction of structures and functional location. All this gives a technical state of the art that could barely have been imagined a few decades ago. But radiology is highly practical. We use images to make inferences about the state of the health of patients and we judge the success of an imaging technique not just on the images themselves but on the radiologist's performance and its effect on patient management [1]. So the success of medical imaging depends on a subjective notion of image quality that is often difficult to define and on factors that influence the ability of the observer to interpret the information [2]. These factors can be summarised into two broad classifications:

1. Those factors that are image dependent and relate to the visual conspicuity of features relevant to the clinical problem; and
2. Those that are image independent; are primarily cognitive in nature and relate to what the observer knows about the visual information in front of him.

We have a wide range of monitoring procedures and tests that help to direct our efforts in the presentation of image features although we rarely know how to define the optimum image for a given diagnostic task. There are many image acquisition, display and processing parameters, and their effects on optimizing images for human interpretation are largely unknown. But we know less still about the cognitive factors influencing perception and allowing the observer to structure the task of interpreting image features; perhaps a better understanding of these factors now deserves our research attention so that we can achieve a better match of image displays to cognitive/perceptual skills. In fact, the image features we currently display for a given examination may change as we know more about the cognitive factors affecting perception and this in turn may lead to changes in image acquisition and processing techniques.

Improving diagnostic performance

How interpretative performance can be enhanced is worth considering. First, the quality of the information and its presentation could be improved. Radiological science has excelled at this, using technical developments to the full, refining and diversifying information and its presentation so that objects critical to the diagnostic question become visible. Second, the taxonomy of those critical objects could be extended and more closely defined. Radiologists have also developed this part of the field extensively and they continue to do so as new imaging modalities offer their contributions. Third, the ability of radiologists to perceive relevant, critical features in the images could be improved.

The work of Birkelo et al in 1947 [3] was the first fully objective evaluation of medical imaging using observer performance data. An editorial accompanying its publication reflected the widespread surprise that variation between readers was greater than the differences between imaging techniques. It had never been considered until that point that expert radiologists might disagree with each other to such an extent on fundamental decisions of diagnostic interpretation. The findings triggered growth in the body of knowledge surrounding perception in medical imaging and the focus was on how and why differences in interpretation (errors) occur [4]. Perceptual and cognitive psychology offered insight into modelling the processes of interpretation, and engineering gave methodologies for evaluating observer performance through the adaptation of signal detection theory to receiver operating characteristics (ROC) studies.

As research in observer performance in radiology has matured so too have developments in ROC methodology thereby enabling more complex experimental designs leading to a better understanding of real-world imaging situations. Experimental work using radiologists as observers has provided significant insight into understanding image interpretation, a task which appears simple but is very complex and one which radiologists execute repeatedly, daily, rapidly and with great skill. Such studies typically concentrate on the errors made in diagnostic imaging and consequently it is easy to forget that radiological error rates are similar to those found in other tasks requiring detailed human visual inspection, thereby demonstrating that such errors often represent the limits of human skill and not necessarily any under-performing issues. It is important to emphasise the vast range of images which radiologists examine and correctly interpret.

Research has typically concentrated upon the causes of false negative errors and demonstrated that three broad

classes can be distinguished: errors due to visual search; detection of radiological features, and their interpretation – broadly demonstrating an increasing cognitive component from the visual registration of the image to the final cognitive decision. Whilst visual search is a necessary part of any image inspection most errors occur due to the other two processes. Experimental studies also illustrate the speed, typically only a few seconds for a static image, with which the inspection process itself is carried out, with additional time then being spent reaching a diagnostic decision.

Consideration of visual search leads to emphasising the importance of foveal inspection of parts of the image. Somewhat contradictorily, and perplexing, is the finding that when an image is presented for very short periods of time – far too short for the eye to move across the image itself – the radiologist's performance is not as poor as may be predicted based on our anatomical knowledge of the eye, and possible abnormalities well outside the foveal area can be identified quite well.

The research over half a century has maintained a clear objective that if we can better understand errors and why they happen, it may be possible to find ways of reducing or eliminating them. The community of investigators committed to this quest is remarkably diverse and shows how complex (and intellectually attractive) the problem is. Publications on the subject typically have author teams from radiology, statistics, physics, radiography, computer science, psychology, vision science, neuroscience, education, mathematics and engineering. Such a broad approach to the problem has ensured that research in the area has always found new roads to explore, but the importance of a strong representative interest from radiology has been the crucial constant in maintaining relevance to medicine in the real world. The importance of that active interest from radiology is particularly important at present because of the changes available in image presentation.

Image displays and manipulation

There is a significant amount of research in developing new displays (*e.g.* liquid crystal displays (LCD)), image processing, image analysis and automated image interpretation (*i.e.* computer-aided detection (CAD)) tools for use in the digital radiology room. Good displays and tools are clearly necessary as the number of images and the complexity of images increase, but what we need to understand is how the radiologist interacts with the displayed information during the reading process in order to determine how we can further improve decision making. We must look at this interaction from a variety of perspectives – perceptual, cognitive and ergonomic. We need to optimize the interface between the radiologists' perceptual/cognitive system and the display, eliminating as many distractions as possible. The radiologist needs to focus on the diagnostic task rather than figure out how to interact with the display just to generate and view a good image.

The visual task faced by radiologists has always altered with time as new imaging techniques have arrived. But the profession is now facing one of the most fundamental changes in a century of practice of interpreting visual information. In moving to an era of filmless image presentation through picture archiving and communication systems

(PACS) there is promise of both help and hindrance. There is no doubt that through processing and manipulating the image at the point of the report there is every opportunity to optimize the image to the visual systems of the observers. But there are many unknowns regarding its best use and how the diagnostic outcomes from workstation reporting compare with hardcopy. These are ergonomic questions and have their solutions in the way that human visual perception and the cognitive activities that ensue relate to workstations as opposed to films and viewing boxes. Current developments in e-science and grid computing presage the advent of the ability to share high volumes of image datasets almost instantaneously across the UK. This means that a radiologist, potentially facing some indecision in classifying a particularly difficult case, could as an aid to their decision making quickly access a range of images of similar appearance, together with their diagnostic outcomes from hospital trusts across the UK. Alternatively, the ability to access and share a great number of images quickly will facilitate new training methods to improve radiological skill nationally. Maintaining high radiological skill levels whilst using technology efficiently and effectively to formulate correct diagnostic decisions quickly is a key issue for the future.

In terms of the display, the exact medium may not be that important as long as its physical characteristics optimize perceptual performance. For years the cathode ray tube (CRT) was the display of choice, but now LCDs and even plasma displays are being used in many departments. Although there are differences between the display technologies, there are a number of important aspects common to them all that are important for optimal medical image perception – luminance, grey scale display calibration or perceptual linearization, and even something as subtle as the type of phosphor used in the faceplate of a CRT display must be considered when characterizing displays for radiology [5]. All of these display characteristics have been shown to not only affect observer performance in terms of decision accuracy, but they also affect visual search efficiency and hence workflow. When eye position is measured as radiologists search various displays, the optimized displays tend to be associated with shorter times to first fixate the lesion of interest during search, shorter times to reach the correct diagnostic decision, shorter overall search times, and fewer redundant comparisons between lesion and non-lesion areas in the images. More efficient search translates into more efficient workflow, increasing productivity and potentially reducing reader fatigue.

The good display is not only a function of the hardware and display medium. The software and user interface are also important. Image processing and image analysis tools should generate images that enhance the information content but should not require prolonged manipulation and processing time. The goal is to provide radiologists' perceptual and cognitive systems with as much useful information as possible without running the risk of overload that could degrade the decision process. The problem is that in many cases these tools are so dependent on users' personal preferences that demonstrating their effect on diagnostic performance is not easy. One area that has shown significant promise towards improving perception and cognition in digital reading is the addition of CAD and diagnosis (CADx) schemes to the armoury of

image analysis tools. [6]. What we still need to understand however, is exactly how these CAD prompts improve performance and what are the optimal ways to provide those prompts. Additionally, it is important that the development and availability of such systems do not detract from the quality and need for radiological skills across the imaging workforce. For instance, initially CAD approaches in breast screening tended to produce considerable numbers of false positive detections as well as potential true detections of abnormality, thereby simply increasing the cognitive load on the radiologist in determining the true from false prompts as well as inspecting the image. Subsequent algorithm developments have greatly reduced the number of erroneous prompts thereby increasing the usefulness of such CAD approaches, however the skill of the radiologist in using such technology remains paramount.

The research community of medical image perception

The Medical Image Perception Society (MIPS) was founded in America in 1997 from its predecessor, the Far West Perception Conference. Its aims are to promote research and education in medical image perception and to provide a forum for discussing perceptual, cognitive and psychophysical issues by radiologists and scientists. For the first time its biannual conference will be held in the UK in 2005, presenting British radiologists and radiological scientists with an opportunity to engage with its work. Research in perception in medical imaging is important because it has enormous potential for improving radiological education, the diagnostic process and for reducing the risk of litigation [7]. Between 30% and 40% of errors in clinical radiology are false negative and thought

to be perceptual; and if one adds to that the smaller error contribution from false positive decisions the total burden from cognitive/perceptual causes approaches 50% of all radiological errors in clinical practice [1, 4]. The entire research effort of CAD and CADx is aimed at catching these errors by modified dual reading or cueing, but it would be a more satisfying outcome if we could fathom the perceptual processes that lead to those errors in the first place; and such an approach has the radiologist at its very centre.

References

1. Kundel H. Reader error, object recognition and visual search. In: Chakraborty D, Eckstein M, editors. Medical imaging 2004. Proceedings of SPIE Vol 5372 (SPIE Bellingham, WA, 2004) 1–9.
2. Manning DJ, Ethell SC, Donovan T. Detection or decision errors? Missed lung cancer from the posteroanterior chest radiograph. Br J Radiol 2004;77:231–5.
3. Birkelo CC, Chamberlain WE, Phelps PS. Tuberculosis case finding. A comparison of the effectiveness of various roentgenographic and photofluorographic methods. JAMA 1947;133:359–66.
4. Krupinski EA. The future of image perception in radiology: synergy between humans and computers. Acad Radiol 2003;10:1–3.
5. Krupinski EA, Roehrig H. Pulmonary nodule detection and visual search: P45 and P104 monochrome versus color monitor displays. Acad Radiol 2002;9:638–45.
6. Huo Z, Giger ML, Vyborny CJ, et al. Breast cancer: effectiveness of computer-aided diagnosis observer study with independent database of mammograms. Radiology 2002;224:560–8.
7. Gunderman R, Williamson K, Fraley R, Steele J. Expertise: implications for radiological education. Acad Radiol 2001;8:1252–6.