The Cognitive Basis for the Design of a Mammography Interpretation Tutor¹

Roger Azevedo, Carnegie Mellon University, Department of Psychology, Pittsburgh, PA 15213 USA

Susanne P. Lajoie, McGill University, Applied Cognitive Science Research Group, 3700 McTavish Street, Montréal, Québéc, Canada, H3A 1Y2

Abstract. The purpose of this paper is to present a cognitively-based and empirically-derived approach for the design of the RadTutor, a prototype computerized tutor to train radiology residents in diagnosing mammograms exhibiting breast diseases. A multitude of computer-based radiology training environments have recently been developed with the objective of supporting the acquisition of radiological expertise. In general, however, these systems have failed in several aspects including a failure to incorporate theoretical perspectives and empirical findings to the design of these systems. This paper outlines the conceptual framework for the development of the prototype which includes: (1) a discussion of the objectives and goals of the radiology residency training program, (2) a review and critique of existing computer-based radiology training environments, (3) a synthesis of an expert-novice study aimed at attaining a cognitive model of problem solving in mammogram interpretation (Azevedo, 1997), (4) a description of the results of analyses of authentic radiology resident teaching rounds, and (5) deriving instructional principles for the design of the mammography tutor.

INTRODUCTION

Radiological expertise is complex, involving several years of acquiring formalized medical knowledge as well as many years of clinical experience. It involves the integration of several distinct bodies of knowledge with separate organizing principles, including physiology, anatomy, pathophysiology, and projective geometry of radiography. Various theoretical frameworks postulate that the attainment of accurate visual diagnostic reasoning abilities involves the interaction of cognitive and perceptual factors.

This paper outlines the cognitively-based and empirically-derived conceptual framework for the development of the RadTutor, a prototype computerized tutoring system to train residents to interpret mammograms. The conceptual framework incorporates the results of a recent cognitive study of mammogram interpretation (Azevedo, 1997) including the cognitive model of mammogram interpretation, the problem solving strategies used by staff radiologists and radiology residents, and the typical case-related errors. Furthermore, the framework is also based on: (1) a critical assessment of the haphazard nature of radiology residency training programs, (2) a review and critique of existing computer-based radiology training environments, (3) an analysis of authentic radiology resident teaching rounds, (4) an assessment of the state of cognitive science and learning in medicine, and (5) instructional principles for the design of the mammography tutor.

¹ An earlier version of this paper was presented at AI-ED97 and published as Azevedo et al. (1997). RadTutor: The theoretical and empirical basis for the design of a mammography interpretation tutor. In B. du Boulay & R. Mizoguchi (Eds.), Frontiers in Artificial Intelligence and Application (pp. 386-393). Amsterdam: IOS Press.
RESIDENCY TRAINING PROGRAM IN RADIOLOGY

The primary goal of the diagnostic radiology training program at McGill University is to produce well-rounded general radiologists who have been exposed to all aspects of subspecialty training (e.g., mammography) and have developed familiarity with all of the imaging techniques (e.g., MRI) and procedures (e.g., endovaginal ultrasound). Becoming a well-rounded general radiologist necessitates a thorough knowledge of the relevant anatomy, physiology and pathology as well as the essentials of medicine and surgery. Radiation physics and biology must be well understood so that the principles can be applied to everyday practice.

Radiology residency programs have traditionally involved an apprenticeship experience, as have other medical education programs. This approach involves centering learning on the patient so that he/she becomes the focal point of all educational, training and research activities (Rubin, 1989). The realistic constraints of the multitude of patients, variability of diseases/disorders, and the complexities of treatments and diagnoses leave little time for additional activities in a busy teaching hospital. The goal of the residency training program is to expose residents to a multitude of pertinent learning, training and research activities aimed at fostering the acquisition of radiological expertise. These activities include annual research days, conferences, formal lectures, and presentations with visiting professors, rounds and reporting sessions. Competence is developed and refined in instructional activities throughout the five-year residency training program. There are however several problems associated with the residency training program, which pose challenges for the design of a computer-based training environment.

In general, radiology residency training is somewhat haphazard. For example, teaching rounds vary in the (1) material taught, (2) quality of presentation by individual radiology staff members, (3) participation of residents during each session, (4) quantity of training experiences, (5) quality of training experiences, and (6) time that staff radiologists have at their disposal for preparing to teach in rounds and lectures. Lastly, the instructional objectives delineated in the residents training manual are too general and thus hamper any attempts at a formalized training program and also limit the objective assessment of residents’ skills. For example, one of the objectives for a 4th year radiology resident is to “have achieved diagnostic and clinical skills at the level of a qualified general radiologist.” This extremely broad objective not only leads to difficulty in assessing these skills but also in terms of designing a computer-based training environment. Overall, this leads to an enormous amount of inconsistency in the training of radiology residents and therefore affects the adequate acquisition of radiological expertise.

COMPUTERIZED TUTORS FOR RADIOLOGY TRAINING

The problem of inconsistency in radiology residency training programs has recently been addressed by the widespread proliferation and dissemination of computer-based training programs (for an extensive review see Lillehaug & Lajoie, 1998). However, most of these systems typically suffer from some major instructional deficits (for an extensive review of these instructional issues refer to Azevedo, Lajoie, Desaulniers, Fleiszer, & Bret, 1997). A multitude of computer-based radiology training environments has recently been developed with the ill-defined objective of supporting the acquisition of radiological expertise. In general, these systems have failed to reach this objective since they: (1) lack a theoretical framework incorporating the empirical evidence from expert-novice differences, cognitive skill acquisition, medical cognition and radiological expertise; (2) are not based on adequate models of learning and instruction; (3) represent technology-driven projects employing various CBI typologies (e.g., linear CBI, branching CBI, multimedia, hypermedia, and WWW); (4) are based on comparison studies of learning effectiveness (between computerized instruction and traditional lectures or several CBI typologies) which lack both methodological and statistical rigor; (5) fail to include the results of authentic analyses of discourse during teaching rounds; (6) adopt behavioral objectives, instructional methods (e.g., linear tutorials), assessment techniques (e.g., multiple choice) and remediation (e.g., canned feedback messages) which are at odds with
recent theoretical and empirical advances in cognitive science, learning, and assessment in illstructured domains; (7) fail to adopt instructional objectives stated in the residency program, and (8) fail to incorporate the computer environments into existing medical curricula. This section presents some examples of more successfully developed radiology training systems, which have addressed the above problems and overcame them.

In recent years there has been a general increase in interest in the application of intelligent tutoring systems (ITSs) in the area of medical training. The rationale for computer-based instruction is based on the assumption that the learner’s cognitive processes can be modeled, traced, and corrected in the context of problem-solving (Anderson, Corbett, Koedinger, & Pelletier, 1995; Greer & McCalla, 1994; Lajoie, 1993, Lajoie, in press; Lajoie & Azevedo, in press; Shute & Psotka, 1994). In recent years, several ITSs have been developed for radiology training.

Examples include the CT Brain Tutor for training radiology residents to diagnose brain tumors from CT scans (Macura, Macura, Toro, Binet, & Trueblood, 1994), and a tutor for training radiology residents to diagnose neurological MRI images (Sharples, duBoulay, Teather, Teather, Jeffrey, & duBoulay, 1995). The extensive work of Sharples and colleagues (Sharples, 1991; Sharples, duBoulay, Teather, Teather, Jeffrey, & duBoulay, 1994; Sharples, Jefferey, Teather, Teather, & du Boulay, 1997) in developing the CT and MRI tutors focuses on accounts of professional practice and skill development and how these issues influence the design of their tutors. They have used statistically-based principles and a structured image description language for teaching radiological image interpretation and the diagnosis of cerebral diseases. Their approach to visual concept tutoring is based on grouping exemplars. Their tutoring approach facilitates the novice to expert transition by assisting the residents in the progression from visual to structural schemas (facilitating rapid pattern matching) and therefore ensuring transfer of skills and learning (Sharples, 1991). Lastly, their tutors aim at training radiologists to view and describe images in a systematic manner.

More recently, Rogers (1995) developed the VIA-RAD tutor based on extensive analyses of verbal protocols obtained from staff and radiology residents during the interpretation of chest x-rays. The tutor is based on the integration of computer-displayed radiological images with cooperative computerized assistance for decision-making. The VIA-RAD system (Visual Interaction Assistant for Radiology) is a blackboard-based architecture, founded on extensive data collection and analysis in the domain of diagnostic radiology, together with cognitive modelling of the interaction between perception and problem-solving. A small prototype of the system has been implemented and tested with radiology professionals.

In sum, these ITSs are based on cognitive science principles of expertise development and incorporate tutoring interventions and tutorial dialogues that are based on analyses of human interactions. As such, an ITS approach would offer consistency and standardization in the training of mammogram interpretation. Therefore, this section presents a cognitively-based and empirically-derived approach for the design of the RadTutor, a computerized tutor for training radiology residents to diagnose breast diseases from mammograms.

THEORETICAL FRAMEWORK: THE NATURE OF RADIOLOGICAL EXPERTISE

Numerous researchers employing disparate theoretical and empirical paradigms have investigated radiological expertise. Three basic “paradigms” that have been applied widely are: (1) search studies which investigate eye movement patterns while experts and novices read x-ray films, (2) signal-detection studies which investigate the ability of novices and experts to detect normal and abnormal film findings, and (3) cognitive research aimed at eliciting the underlying cognitive and perceptual factors involved in radiological expertise. Relatively few studies (Azevedo, 1997; Faremo, 1997; Lesgold, Rubinson, Feltovich, Glaser, Klopfer, & Wang, 1988; Rogers, 1992; 1996) have actually investigated the underlying cognitive and perceptual factors involved in radiological diagnosis. As a result, a fundamental understanding of the constitution and acquisition of expertise in other radiological sub-specialties such as
mammography has yet to be determined. The following section presents a brief overview of a recent study in mammogram interpretation and presents the major findings.

**A Cognitive Study of Mammogram Interpretation**

A recent study by Azevedo (1997) examined the problem solving strategies used by staff radiologists and radiology residents during the interpretation of difficult mammograms. Ten radiologists and ten residents diagnosed 10 cases under two experimental conditions (authentic and augmented). In the authentic condition, standard unmarked mammograms were used. Mammographic findings were highlighted on a second set of the same cases for the augmented condition. Verbal protocols were analyzed and revealed that mammography interpretation was characterized by a predominant use of data-driven or mixed-strategies depending on case typicality and clinical experience. Repeated measures ANOVAs revealed that the radiologists scanned the cases significantly faster than the residents. No group differences were found in the number of radiological findings, radiological observations, and number of diagnoses across experimental conditions. Frequency analyses revealed that regardless of experimental condition both groups (1) used the same types of operators, control processes, diagnostic plans, (2) committed the same number of errors, and (3) committed case-dependent errors. Overall, the fact that few differences were found between the groups on the various measures may be due to the fact that mammogram interpretation is a well-constrained visual cognitive task.

**Implications for the Design of the RadTutor**

The results of this study have served as one source of empirical basis for the design of the RadTutor. The following is a brief discussion as to how the results have been incorporated into the RadTutor. The content analyses of the areas of breast disease and mammography have been used to construct the domain knowledge module of the prototype as a series of production rules. The cognitive task analyses (CTAs) based on extensive interviews with the domain expert and a surgeon specializing in breast diseases were used to develop the overall instructional sequencing for each case and to build the system’s expert module. Additional sources of support were derived from other CTAs that were conducted with various physicians as part of a multimedia project on the management of breast diseases (Rahilly, Saroyan, Greer, Lajoie, Breuleux, Azevedo, & Fleiszer, 1996).

The cognitive model characterizing mammogram interpretation consisting of 7 steps served as the overall instructional sequencing for the system. These steps include: (1) reading a clinical history, (2) placing a set of mammograms on a viewbox and identifying individual mammograms in the set, (3) visually inspecting each of the mammograms either with or without the use of a magnifying glass, (4) identifying mammographic findings and observations, (5) characterizing mammographic findings and observations, (6) providing a definitive diagnosis or a set of differential diagnoses, and (7) specifying subsequent examinations. In addition, the system is capable of determining if the user is employing a data-driven and/or a mixed problem solving strategy. These steps are used as a generic model of diagnostic problem solving in mammogram interpretation. Data-driven involves solving a problem from data to diagnosis. For example, a subject read the clinical history, scanned the mammograms, characterized the mammographic findings and provided a diagnosis. In contrast, the mixed strategy involved a combination of data-driven and goal-driven problem solving strategies. For example, a subject read the clinical history, scanned the mammograms, provided a set of differential diagnoses and then proceeded to inspect the mammograms for specific findings and reduced the number of differential diagnoses. The system monitors the evolution of the user’s problem solving behavior (during the resolution of a case) and predicts if he/she is engaged in one or the other problem solving strategies. This aspect of the prototype is extremely critical in identifying errors and providing the appropriate level of scaffolding.

The problem solving operators used by both staff and residents indicated that both groups made extensive use (76% of the time) of 4 particular operators (data examination, data acquisition, data exploration, and hypothesis generation). Data examination involved three
operators which were employed to interpret cues presented in a set of mammograms. Generally, these operators were used to determine the significance of a cue in terms of whether it represents a radiological observation or radiological finding. Once a cue was recognized and interpreted, its significance was used as an input for subsequent information-processing operations. The third operator is used to determine if a mammographic cue is an actual mammographic cue or a film artifact. 

**Data acquisition** involved two operators which were used to acquire clinical cues from the clinical history and set of mammograms. The subject typically read the clinical history out loud and inspected the mammograms while placing them on a view box. 

**Data exploration** involved operators which were employed to conduct a more detailed review or further interpretation of cues previously selected from the clinical history and set of mammograms. Typically, it involved further characterizing mammographic cues (observation and findings) according to their location, shape, size, margins, density, number, distribution, description, and/or associated findings, it may involve noting mammographic cues that were not present on the films, and re-reading cues from the clinical history. 

**Hypothesis generation** involved operators which were used to generate diagnostic hypotheses. Specific operators provide different means for generating hypotheses in response to a particular data cue, pattern of cues, data classification, or other diagnostic hypotheses. Is it important to note that cases with multiple “diagnosable” benign and malignant findings required the use of separate operators for each mammographic cue. This information was used primarily in the design of the different levels of instructional scaffolding and interface. For example, data acquisition, examination, and exploration meant that the interface was built to display the case history (data acquisition) and set of mammograms (data acquisition) and allow the user to manipulate the images for better feature characterization or comparison (data exploration). Similarly, the system provides extensive instructional scaffolding during the hypothesis generation phase to ensure that the user has proposed the appropriate hypothesis level (e.g., malignant versus infiltrating ductal carcinoma).

The verbal protocols analyses indicated that diagnostic planning (i.e., propose further medical examinations) was the most frequent control process used by the subjects. As such the interface was built so as to allow the user to list more than one medical examination. This aspect of the prototype is associated with an extensive discussion of the benefits associated with each subsequent examination.

The error analyses revealed five types of errors including:

1. a **perceptual detection error** (failure to detect a mammographic finding),
2. a **finding mischaracterization error** (incorrect characterization of a mammographic finding),
3. a **no diagnosis error** (detection, correct identification, and characterization of a mammographic finding but a failure to make a diagnosis)
4. a **wrong diagnosis error** (detection, correct identification, and characterization of a mammographic finding but proposing a wrong diagnosis), and
5. a **wrong recommendation error** (correct detection and characterization of a mammographic finding, and proposing a diagnosis at some level of abstraction, but proposing an inappropriate subsequent examination).

The results suggest that regardless of level of expertise, the type of error committed is case-related. The results suggest that the clinical history, and more importantly, the mammographic manifestations are critical in determining the types of errors committed by radiology personnel. Each type of error is presently being formalized as a production rule and is integrated in the expert module. Furthermore, each error type is also associated with a specific instructional scaffolding strategy. For example, a finding mischaracterization error is associated with an instructional strategy that focuses the user’s attention on the part of a mammographic finding which was mischaracterized (e.g., the border of a mass).

The process of identifying error commission is facilitated by the fact that the analyses indicated errors to be case-dependent. For example, cases with atypical mammographic
manifestations are highly likely to produce a finding mischaracterization error. In sum, the empirical results of Azevedo’s (1997) study including the cognitive model of mammogram interpretation, problem solving strategies, problem solving operators, control processes and error analyses have served as one source of empirical basis for the design of the RadTutor.

ANALYSES OF THE RADIOLOGY TEACHING ROUNDS: ELICITING TUTORING STRATEGIES

The tutoring strategies and levels of instructional scaffolding adopted in the prototype are also based on the authentic analyses of radiology teaching rounds (Azevedo, Lajoie, Desaulniers, & Bret, 1996). These analyses examined the diagnostic problem solving processes and teaching methodologies employed by a staff radiologist teaching six residents during two one-hour mammography rounds. However, this section focuses specifically on the discourse analyses of the teaching methods used by the staff radiologist during the mammography teaching rounds.

The topographical view of a typical teaching round in mammography is illustrated in Figure 1. A round is typically comprised of an expert (staff radiologist in charge of the teaching round), residents (denoted by the circles – numbers indicate residency level), a resident who is solving a case (e.g., R3), and a set of mammograms displayed on the viewbox (including the mediolateral [MLO] and cephalocaudal [CC] views of the left and right breasts). The two arrows represent the directionality of the discourse that was analyzed. The first arrow indicates R3 solving the case that is presented on the viewbox. The second bi-directional arrow indicates (1) pedagogical strategies that the staff radiologist used to support the resident’s diagnostic problem solving and (2) the resident’s request for various levels of support from the staff radiologist. The results of the analyses are presented below.

![Figure 1. A topographical view of a typical radiology teaching round.](image_url)
redirecting their viewing process, and subsequently fading all instructional support when the resident demonstrated the capability to pursue the task on his/her own. In the RadTutor, coaching and hints are provided through text messages, pop-up text messages, and highlighting of mammographic findings and observations. In addition, multiple levels of instructional scaffolding have also been delineated based on the results of the dissertation and interviews with the expert staff radiologist. For example, there are three levels of instructional scaffolding for the process of highlighting mammographic findings. These include (1) a global strategy to indicate that a mammographic feature has been highlighted on the wrong breast, (2) an intermediate strategy to indicate that a mammographic feature has been highlighted on the correct breast but on the wrong mammographic view (e.g., CC instead of MLO), and (3) a local strategy to indicate the exact location of the mammographic finding.

The finest illustrations of articulation were observed when the radiologist externalized her reasoning process beginning with the assignment of probabilities to pathological features, followed by the systematic elimination of competing differential diagnoses until the definitive diagnosis was achieved. This teaching method was especially valuable since all residents may potentially benefit from the externalization of the expert's diagnostic problem solving. In the RadTutor, articulation is being implemented by using digitized video clips of the expert staff radiologist diagnosing a case while the tutoring system highlights the mammographic findings and observations on the digitized mammograms.

In summary, this section has presented a cognitively-based and empirically-derived approach for the design of the RadTutor. The design approach is based on Azevedo's (1997) study of problem solving in mammogram interpretation, and the tutoring strategies elicited during radiology teaching rounds.

**BRIDGING THE GAP: DERIVING INSTRUCTIONAL PRINCIPLES FROM COGNITIVE SCIENCE AND EMPIRICAL STUDIES**

The following section delineates instructional principles derived by integrating the empirical research on cognitive science, medical cognition, medical education, radiological expertise, mammogram interpretation, and analyses of radiology rounds. The successful integration of these sources of evidence are critical to the design of a theoretically-based and empirically-derived prototype for training both residents and staff radiologists in the interpretation of mammograms. Each of the four instructional principles posited is supported by existing theoretical and empirical evidence described previously in this paper. The instructional principles incorporated in the mammogram interpretation prototype include the principles of (1) multiplicity, (2) activeness, (3) accommodation and adaptation, and (4) authenticity (for an extensive overview of some of these principles for supporting computer-supported problem-based learning see Koschmann, Kelson, Feltovich, & Barrows, 1996).

The **principle of multiplicity** is based on the concept that knowledge is complex, context-sensitive, inter-related and thus instruction should promote multiple perspectives, representations and strategies. This principle is based on the theory of cognitive flexibility (Spiro, Feltovich, Jacobson, & Coulson, 1991) in medicine that emphasizes the use of multiple knowledge representations and repeated exposure to instructional content. According to this principle, single mental representations and unitary learning approaches are insufficient for capturing the nature of complex instructional materials and knowledge application in ill-structured domains (such as radiology). A recent study examining the effectiveness of hypermedia versus traditional lectures in radiology has demonstrated the use of multiple knowledge representations through the use of text, digitized video clips and animations of radiology physics concepts to facilitate learning and knowledge application (Shaw, Azevedo, & Bret, 1995). The RadTutor provides the resident with a stock of breast disease cases that can be accessed in a structured manner according to diagnostic categories, specific mammographic manifestations (i.e., findings and observations), and relevant clinical history cues.

The **principle of activeness** is based on the concept that learning is an active process, requiring mental construction and manipulation of the subsymbolic (e.g., gray-scale densities
exhibited on mammograms) and symbolic representations (e.g., clinical findings exhibited on mammograms and relevant clinical history findings) that comprise the task environment. Therefore, instruction fosters knowledge construction through problem-solving activities that lead to the development of skill acquisition. This principle reflects the nature of learning through active construction of knowledge facilitated by problem-solving activities. Effective instructional methods should promote planning, reasoning, goal-directed problem-solving, and reflection. This principle reflects the empirical findings in the areas of cognitive skill acquisition (VanLehn, 1996) and the development of expertise (Ericsson & Charness, 1997; Ericsson & Lehmann, 1996). In the RadTutor, instruction fosters knowledge construction through meaningful problem-solving activities that facilitate skill acquisition and the development of expertise.

The principle of accommodation and adaptation is based on the concept that the learning process is to a large degree affected by the extent of the learner's existing knowledge. As such, instruction facilitates adaptability by building upon the learner's existing knowledge, monitoring learner progress and rectifying misconceptions when they arise, and fostering the development of metacognitive skills. A rule-based domain knowledge module and a student modelling approach are presently being considered based on the well-constrained nature of the domain of mammography.

The principle of authenticity is based on the concept that learning is sensitive to contextual factors, which determine the usability of what is learned, and the extent of skill transfer. Therefore, instruction should provide learning activities that are required in the domain, that are valued in the real-world context, and that emulate the real-world environment as much as possible. This principle reflects the recent claims by advocates of situated cognition (Clancey, 1997; Greeno, 1998) regarding the need to study the contextual and situational aspects of the cognitive phenomena being studied. In the RadTutor, the problem-solving activities resemble what is routinely encountered in a resident's work environment. For example, it provides the tools typically used to solve mammogram cases (such as a magnifying glass and a ruler to measure masses and lesions).

**DESCRIPTION OF THE RADTUTOR PROTOTYPE**

The RadTutor prototype offers instructional approaches such as observation, modelling, coaching, scaffolding, fading of assistance, shared problem solving between tutor and student, and situated learning in the context of subsequent knowledge use. Residents are engaged in explicitly justifying hypotheses based on the evidence gathered and rating the medical evidence. The prototype traces the user's problem-solving activities and provides microadaptive and macroadaptive interventions based on the (1) collection of pertinent data by observation, (2) selection of medical evidence, (3) correctness in interpreting the medical evidence, (4) construction of a hypothesis based on the integration of the data, and (5) diagnostic accuracy. The tutoring interventions are based on (1) the concept of creating an effective problem space (EPS) reifying the empirical results of the Azevedo (1997) study, (2) cognitive task analyses obtained from consulting radiologists and residents having varying levels of expertise, and (3) the analyses of authentic radiology rounds between staff radiologists and residents with varying levels of expertise. Residents are forced to select their hypotheses from a list of options (e.g., carcinoma, fibroadenoma, etc.) and estimate their level of confidence in those hypotheses. The system monitors all of these actions and determine what advice, scaffolding and hints to generate in response to the user's problem-solving activities. Hyperlinks are presently being incorporated into the system allowing the user access to detailed explanations regarding pertinent aspects of each case. Finally, an important component of the prototype is it’s ability to provide the user with a graphical representation of the expert's reasoning process in solving each case (using digitized video). The ITS architecture will have hypermedia capabilities such as the hyperlinks, allow access to digitized video clips and audio segments, and display digitized radiological assets (e.g., mammograms) retrieved from CD technology.
The following section illustrates the RadTutor’s overall framework. It also provides a detailed description of each of the system’s instructional features including (1) present the patient history, (2) show a set of mammograms, (3) identify mammographic observations, (4) characterize the mammographic observations, (5) identify the mammographic findings depicted on the each mammogram, (6) characterize mammographic findings on each of the mammograms, (7) provide a diagnosis (for each findings, if required), and (8) specify a subsequent examination for a mammographic finding (if required).

Figure 2. The interface of the mammography prototype.

Figure 2 presents the interface of the mammography prototype. The interface is divided into seven separate components: (1) a number of pull-down menus allowing the user to select several actions (e.g., circle a region, select a region, delete region, rectify an action, make a diagnosis, and propose subsequent examinations), (2) a clinical history box, (3) the mammogram display area, (4) the tutor’s dialogue box, (5) the resident’s dialogue box, (6) a list of mammogram observations, findings, and diagnoses (the actual content depends on the instructional sequencing), and (6) a differential diagnosis box.

Solving a typical breast disease case would involve the following steps. The tutor provides the clinical history (the case of a 60 year old woman presented with a mass in the right breast). The tutor then displays the mammograms in a random order whereby a resident can re-position the images in the mammogram display area. The tutor then requires the resident if there are any significant observations to be highlighted. Depending on the resident’s response and its correctness, the tutor will either (1) identify and characterize the observations, or (2) bypass these two steps and go directly to the identification of mammographic findings. During the identification of observation step the resident has to highlight the observation on the mammogram and if correct the tutor characterizes the observation by displaying a text label and arrow pointing to the area. If the region is incorrectly highlighted as an observation then the tutor rehighlights the area and then displays a text label and arrow pointing to the area. If there are multiple observations the tutor will then ask the resident if more observations are exhibited on the mammograms and the cycle is repeated until all observation are identified and characterized.
The tutor instructs the resident to circle and identify the critical findings and to select a diagnosis. The tutor provides the resident with feedback concerning the placement of the mammograms and prompts the resident to look for findings. The resident then has the option to either select the critical findings from a list of features or go directly to the images and highlight the critical features. For example, if the resident selected large fibroadenoma from the feature identification list. The tutor will respond with a simple feedback message stating that the selected finding is critical to the diagnosis but incorrectly labelled. The resident may have selected a diagnosis of benign and placed it in the active diagnosis box. The resident may highlight a partly well-circumscribed mass in the right breast with the aid of the tools provided in the floating palette. He/she may also characterize the mammographic finding (I think the mass is well-circumscribed). The tutor will subsequently intervene by asking the resident to rectify his characterization since the mass is partly well-circumscribed and therefore should yield a correct diagnosis of suspicious for malignancy. In mammography, diagnoses may be stated at different levels of abstraction, however, the level of specification is not always critical. For example, diagnosing a case as malignant or medullary carcinoma is equivalent since the subsequent examination would entail performing a biopsy. Furthermore, during the problem solving process diagnoses may be either refined (e.g., from an initial diagnosis of benign to a final diagnosis of hamartoma) or generalized (e.g., from in situ carcinoma to malignant).

The resident can select as many features from the features identification list and can highlight as many mammographic regions as he/she wants. At a micro level, tutoring interventions, instructional scaffolding, and resident queries are based on one of the features identified. However, at the global level, the scaffolding and tutoring interventions are based on the number of findings selected and highlighted, the accuracy of their characterization and differential diagnoses selected. The dialogue between tutor and resident is presented in dialogue boxes and the termination of a dialogue sequence is represented by the dotted lines.

- **Show Patient History**

  First, the user is presented with the patient history. For example, a 60 year old woman presented with a mass in the right breast.

- **Show The Set of Mammograms**

  Secondly, the user is presented with the patient’s mammograms (typically comprised of left and right mediolateral oblique and craniocaudal views). The environment allows the user to place the mammograms in any order and position by selecting and dragging the image.

- **List and Highlight all of the Findings Depicted on each Mammogram**

  Thirdly, the user traces the findings exhibited on each mammogram and/or selects them from a long list of possible findings. For example, the critical findings related to case #1 include: dense fibroglandular tissue bilaterally, large mass in the right retroaereolar area, mass is partly well-circumscribed, mass is not well-circumscribed. If the findings identified are correct, confirmatory feedback will be provided and the findings will be highlighted on each of the mammograms. However, if the findings listed are incorrect the tutor will provide different levels of scaffolding for each depending on the correctness and type of error. The levels of instructional scaffolding and content of feedback messages are based on the results of the authentic analyses of the teaching rounds and CTAs conducted with the consulting staff radiologist and several other sources as mentioned earlier in the paper. At the macro level, the tutor scaffolds learning by asking a question regarding the region of interest (ROI) where the finding may be located. At the micro level, the tutor asks a specific question (such as “Is there fibroglandular tissue in the upper outer quadrant of the left breast?”). After three attempts the tutor highlights the finding on the mammogram. The use of feedback messages and hints are based on a recent meta-analysis on the effects of feedback in computer-based instruction (Azevedo & Bernard; 1995a; 1995b). This cycle continues until the resident identifies and lists all of the findings exhibited in each case.
• **What is the Diagnosis?**

Once a resident identifies and selects all of the relevant findings, he/she is asked to provide a diagnosis. If the user provides the correct diagnosis he/she will move on to the next step. For example, the correct diagnosis for case #1 is a *mucin-producing carcinoma* (colloid carcinoma or mucinous carcinoma). However, if the user provides an incorrect diagnosis (e.g., *fibroadenoma, cyst, metastatic disease*), different levels of scaffolding are available depending on the correctness of the diagnosis. The first level of intervention consists of reviewing the patient's history and the critical mammography findings. This level of scaffolding is aimed at establishing a diagnostic schema by reviewing the results of the previous step (List and Highlight all of the Findings Depicted on each Mammogram).

The second level of intervention displays the digitized clip illustrating the consulting radiologist's "externalization" of the clinical reasoning involved in correctly diagnosing the mammograms. This approach includes an extensive elaboration of the radiologist’s case resolution by establishing the critical findings on the mammogram, selecting the findings from the list of possible findings, and then reasoning the case through. Presently, we are incorporating other features into the tutor such as (1) a glossary of breast disease terms, (2) digitized images and video clips demonstrating how to place mammograms on a viewbox, (3) a scratch pad so that residents can “cut and paste” screens from the tutor and either send them to print or save them to disk, (4) a list of radiology and medical references including abstracts, and (5) network access to other teaching files and on-line resources such as electronic radiology journals.

• **Are There Any Subsequent Examinations?**

This last step is most critical in malignant cases since they require subsequent medical examinations such as a biopsy, ultrasound, fine needle and/or aspiration. As such the user must specify the type of subsequent examination required for a specific case. For benign cases the answer is always "none". Furthermore, the system provides different levels of instructional scaffolding depending on the correctness of the examination provided.

**CONCLUSION**

This paper outlined a cognitively-based and empirically-derived conceptual framework for the development of the RadTutor, a prototype computerized tutor to train radiology residents in diagnosing mammograms. The conceptual framework is based on: (1) a critical assessment of the haphazard nature of radiology residency training programs, (2) a review and critique of existing computer-based radiology training environments, (3) a recent cognitive study of mammogram interpretation (including the cognitive model of mammogram interpretation, the problem solving strategies used by staff radiologists and radiology residents, and the typical case-related errors), (4) an analysis of authentic radiology resident teaching rounds, and (5) instructional principles for the design of the mammography tutor.

The future of the RadTutor is being pursued in several directions including theoretical, empirical, and tutor development. First, an expert-novice study is planned to investigate the interaction between perceptual and cognitive factors in mammogram interpretation by converging multiple sources of data including verbal protocols, eye-movements, detection ability, and underlying knowledge structures. The results of this study will enhance our understanding of the diagnostic process in mammogram interpretation. Second, an evaluation of the existing prototype with radiology residents associated with McGill University’s and the University of Pittsburgh Medical Center’s teaching hospitals is scheduled to commence in the near future. Several aspects of the RadTutor will be modified based on the result of the empirical results of the expert-novice study and the evaluation of the prototype’s effectiveness in supporting the diagnostic problem solving behavior of radiology residents. Lastly, Anderson’s (1993) ACT-R theory of cognitive skills is presently being used to model the
The Cognitive Basis for the Design of a Mammography Interpretation Tutor

cognitive aspects of mammogram interpretation and will also be used to develop the RadTutor’s model-tracing student model.

Acknowledgments

The design and development of the RadTutor prototype is based on the results of the first author’s Ph.D. thesis and was prepared for publication while the author was working at the Department of Psychology of Carnegie Mellon University as a postdoctoral fellow funded by the Social Sciences and Humanities Research Council of Canada (SSHRC). This project was funded by Dr. David M. Fleiszer and McGill University’s Medical Informatics Committee (Molson Project) and partial funds from a doctoral fellowship from the SSHRC also awarded to the first author. Partial funding was also provided by the Quebéc government’s Ministry of Industry, Science, Commerce and Technology (MISCT) and the SAFARI project. The authors would like to thank Xiaoyan Zhao for the development of the RadTutor prototype, Drs. Desaulniers and Fleiszer for their expertise in the areas of mammography and breast disease, Sonia Faremo for insightful comments and various discussions on the topics presented in this paper, and Daniel Leibu and Hugo Dufort for the initial development of the prototype.

References

Azevedo and Lajoie


