Perceptual Skill, Radiology Expertise, and Visual Test Performance with NINA and WALDO[1]

Calvin F. Nodine, PhD, Elizabeth A. Krupinski, PhD

Rationale and Objectives. The goal of this study was to determine if radiologists possess superior visual search and analysis skills compared with those of laypeople.

Materials and Methods. In two experiments, radiologists and laypeople searched one of two complex pictorial scenes for hidden targets. Eye position was recorded during the search. Two measures of performance were obtained: accuracy of detecting targets as measured by using alternative free response receiver operating characteristic analysis and visual search efficiency as measured by using eye position analysis.

Results. There were no statistically significant differences in detection performance between radiologists and laypeople for either of the search tasks. Radiologists took longer on average to search the images and to first fixate on the targets than did the laypeople. For both groups, true-positive and false-positive decisions were associated with longer dwell times than true-negative decisions. As with radiology search tasks, false-negative decisions were also associated with longer dwell times than true-negative decisions.

Conclusion. Performance on two visual search and detection tasks indicate that radiologists do not possess superior visual skills compared with laypeople. Radiology expertise is more likely to be a combination of specific visual and cognitive skills derived from medical training and experience in detecting and determining the diagnostic importance of radiographic findings.

Key Word. Diagnostic radiology, observer performance.

There is a common assumption that radiologists are better visual analyzers than most of their medical colleagues. Whether this visual skill is innate or acquired has been the subject of numerous studies (1–5). Generally, results from perceptual tests tend to correlate fairly well with general ratings of diagnostic abilities (2–5) but less well with results of specific diagnostic tasks such as pulmonary nodule detection (1). Thus, the answer to this question is unfortunately not easy to determine, primarily because innate visual skills quickly become contaminated by training and experience (6–8). Furthermore, visual testers have generally assumed that the radiologist's task is largely a visual one.

There is also a great deal of cognitive interpretation that goes into the reading of an x-ray image. For example, in addition to searching for abnormalities, radiologists "read" medical images for anatomic and pathologic content as they search the image. This point is largely overlooked by researchers who have developed visual tests. The radiologist's report typically contains a description of the findings resulting from the search and an interpretation of the findings considered in the context of the patient's history. This separation of description from interpretation in the report provides radiologists with a framework for carrying out visual and cognitive aspects of the diagnostic radiology task in much the same way as instructions provide observers with a framework for car-

Acad Radiol 1998; 5:603-612

[1] From the Department of Radiology, University of Pennsylvania, 3400 Spruce St, Philadelphia, PA 19104-6086 (C.F.N.); and the Department of Radiology, University of Arizona, Tucson (E.A.K.). Received January 26, 1998; revision requested March 19; revision received April 9; accepted April 15. Supported in part by US Army MRMC DAMD17-97-1-7130. **Address reprint requests to** C.F.N.

rying out an experimental test in the laboratory.

Artist Al Hirschfeld has been hiding the word "NINA" (his daughter's name) in line drawings of theatrical scenes that have appeared in the *New York Times* for over 50 years. The hide-and-seek game of finding the name NINA in Hirschfeld's drawings illustrates basic perceptual principles of detection, discrimination, and decision making that are commonly encountered in radiology search tasks. Hirschfeld's hiding of NINA is typically accomplished by camouflaging the letters of the name and blending them into scenic background details, such as wisps of hair and folds of clothing. In a similar way, pulmonary nodules and breast lesions are camouflaged on radiographs by anatomic features of the chest or breast. Hirschfeld's hidden NINAs are sometimes missed because they are perceptually integrated into a gestalt overview of the picture, rather than differentiated from background features during focal scanning. This may be similar to the overlooking of an obvious nodule behind the heart in a chest radiograph. Because it is a search game, Hirschfeld assigns a number to each drawing to indicate how many NINAs he has hidden so as not to frustrate his viewers. In the radiologists' task, the number of targets detected in a medical image is presumed to be determined by combining perceptual input with probabilities generated from clinical history and viewing experience. Thus, in the absence of truth, searching for abnormalities in radiographic images creates opportunities for recognition and decision errors (eg, false-positive and false-negative decisions).

Reading of medical images requires both search and interpretation of radiologic findings within an anatomic image context. The task of searching, interpreting, and reading the medical image combines perceptual and cognitive skills that most test developers have failed to appreciate. We have found experimental evidence indicating that observers have difficulty carrying out both visual search and interpretation tasks simultaneously in a testing situation. For example, in one study when observers were instructed to search for NINA, they had mixed success finding the target (9). Afterward, the observers were asked to describe the scenes they had just searched. They could not describe the gist of the scene nor identify the main characters, even though the charcters were familiar well-known actors who they later identified when shown the drawings. Maybe this is why radiologists typically dictate the report while looking at the x-ray image, the implication being that the search has revealed findings and the image is used as a reference map during the generation of the report that both describes and interprets the findings.

The above discussion points out the need to analyze and identify task requirements before selecting tests to measure and compare radiologists' performance skills. It is clear that visual search skill is one component of the radiologist's task. Others include the ability to (*a*) disembed figures from background as in hidden figures tests, (*b*) form an instantaneous gestalt or global interpretation of a scene to obtain the gist and identify regions of interest for search, (*c*) extract distinctive features that signal perturbations in anatomic image scenery, and (*d*) weight the importance of distinctive features extracted from visual input during the search with hypotheses generated from experience in diagnostic decision making.

We looked for a test that taps these skills. This article reports the results of our experiments in which we used two visual search tasks that come close to meeting the task requirements in radiology, as listed above. We compared the performance of radiologists with that of laypeople in searching art pictures to find hidden targets. The art pictures do not presuppose any prior knowledge in searching for a target. This is a way of equating observers for experience. The targets were the word "NINA" embedded in Al Hirschfeld's line drawings of theatrical scenes (10) and color drawings by Handford of the character WALDO embedded in people-cluttered scenic backgrounds (11). As with the anatomic scenery in radiographs, the artistically represented scenery in our test pictures typically acts to camouflage the target, and thus the art-picture search tasks have some of the same characteristics as the radiographic search task. In addition, both the medical-image and the art-test targets have distinctive features that provide a perceptual basis for visual differentiation of target from background. Finally, detection and recognition of targets in both medical images and test pictures are sufficiently ambiguous that observers can effectively provide confidence ratings for their decisions. Thus, we used standard detection measures to evaluate the test results.

To summarize, this article examines the types of visual skills that are useful to radiologists and how training and experience influences these visual skills. We present data from two studies in which the visual skills of radiologists were compared with those of laypeople on visual search tasks in which *both* groups were inexperienced. In each case, the subjects were required to search a picture and find a hidden target. This task is not unlike searching a chest x-ray image for a lung nodule or mammogram for a breast lesion.

MATERIALS AND METHODS

Two types of picture search tasks were used: line drawings by artist Al Hirschfeld in which the target was the word NINA embedded in the line drawing and color drawings by the artist Martin Handford in which the target figure was a character named WALDO embedded among numerous colored line figures. Radiologists and laypeople were recruited as observers from the University of Pennsylvania and the University of Arizona Medical Center. Five radiologists and six laypeople from Pennsylvania served as observers for the NINA test. Seven radiologists and seven laypeople from Arizona served as observers for the WALDO test.

NINA Test

Each observer was given a test booklet that contained photocopies of 42 Hirschfeld drawings from *The World of Hirschfeld* (10). Each drawing contained zero to seven hidden NINAs (average, two per picture). After an introduction and illustration of the NINA search task, observers were paced through the test booklet at the rate of 60 seconds per picture to find and circle the word NINA and rate their confidence in detection. A beeper sounded 10 seconds before the time limit so that observers could indicate any remaining uncircled NINAs and turn the page to a new picture. When a NINA was detected and circled, observers were asked to rate their confidence in interpreting the line configuration as a NINA (3, definite; 2, probably; 1, maybe). The number below Hirschfeld's signature that indicated how many NINAs were hidden in the picture was removed so that observers did not know how many NINAs to search for. At an average viewing distance of 40 cm, each 21.6 × 28.0-cm picture page subtended a visual angle of approximately 28°. The NINA targets ranged in size from 0.7 cm (less than 1°) to 6 cm (more than 8°). A chest x-ray image viewed at the same distance subtends a visual angle of approximately 42° and a 1-cm nodule subtends 1.4°. Eye position was monitored for a subset of three NINA pictures that were viewed by 10 observers. The observers had little or no experience with Hirschfeld's NINA drawings.

WALDO Test

After an introduction to and illustration of the WALDO search task, observers were shown a set of 10 full-size, 48 × 31-cm color poster pictures from *Where's WALDO? The Magnificent Poster Book* (11). Each picture contained one WALDO plus some foils: Wilma, Wenda, Odlaw, and numerous other characters typically reported as WALDO (ie, false-positive findings). The foils ranged in size from 0.5 × 0.3 cm (1°) to 1.8 × 0.5 cm (3.4°). Observers were given up to 2 minutes to find and point out WALDO. Feedback was given by the experimenter (E.A.K.) as observers searched for WALDO. False-positive sightings were noted as such to the observers, and the observers were told to continue searching for the real WALDO. All observers knew who WALDO was, what he looked like, what color his clothes were, and the fact that WALDO was often partly obscured by other people or things in the picture. Confidence ratings were obtained when WALDO was detected; however, because observers did not use the entire scale, the confidence rating data were discarded. The full-color poster pictures subtended a visual angle of approximately 46° at a 30-cm viewing distance, and the WALDO targets ranged in size from 0.7 cm (1.3°) to 2.3 cm (4.4°).

Data Analysis

Two measures of performance were obtained: accuracy of detecting targets and visual search efficiency as measured by using eye position analysis. Alternative free response receiver operating characteristic (AFROC) analysis (12) was used for the NINA task because the pictures typically contained more than one target. The area under the AFROC curve, A1, is the estimated probability of any given true target being rated higher than the most suspicious nontarget on the same image. For the NINA study, A1 was estimated from the highest rated correctly localized true-positive responses relative to the highest rated false-positive response per picture. For the WALDO study, the observers used the rating value 6 (definitely WALDO) when they found a WALDO (a true-positive response) or a WALDO look-alike (false-positive response). Observers were always convinced that they had definitely found WALDO even when they were wrong! Because of this, the probability of a correct first-choice localization could not be estimated. Therefore, A1 was estimated from the probability of the first correctly localized true-positive response on WALDO relative to all prior false-positive responses per picture.

Analysis of eye position data focused on three measures of search efficiency: search time to fixate on the target, total viewing time, and cumulative gaze duration (visual dwell time). The eye position data for NINA testing was limited to a subset of 10 observers (four radiologists and six laypeople) and three pictures. Four records were lost due to poor calibration, for a total of 26 records.

Table 1
AFROC A1 Area Values in NINA and Estimated A1 Area Values for the WALDO Test Pictures

	NINA		WALDO	
Observer No.	Radiologists	Laypeople	Radiologists	Laypeople
1.	.526	.566	.600	.900
2	.511	.728	.683	.750
3	.482	.528	.650	.650
4	.552	.874	.683	.550
5	.772	.639	.650	.600
6	ND	.802	.500	.733
7	ND	ND	.783	.650
Mean	.569	.689	.650	.690
SD	.136	.116	.086	.116

Note.—ND = no data, SD = standard deviation.

Table 2
Search Time to First Fixate on the Target in NINA and WALDO Test Pictures

	NINA		WALDO	
Search Time	Radiologists	Laypeople	Radiologists	Laypeople
Mean	16.20 (20)	9.99 (35)	26.24 (70)	22.44 (70)
SD	8.03	8.62	22.93	19.68

Note.—Search time data are given in seconds. Numbers in parentheses are number of records. SD = standard deviation.

The eye position data from the WALDO test consisted of 140 records from seven radiologists and seven laypeople, each of whom searched 10 pictures. The three measures of search efficiency were analyzed by using *t* tests and analyses of variance.

A 4000SU Eye-Tracker (Applied Science Laboratories, Bedford, Mass), which records pupil and corneal reflections with an infrared reflection source, was used to record eye position in both studies. The system is accurate to within 1°. For initial calibration purposes, observers were seated in front of the display and the observer's head was stabilized in a chin rest. After initial calibration, the chin rest was removed and the observer was allowed to change position if desired. The 4000SU system comes with a head tracker so that observer head motion is recorded and integrated to adjust for eye position changes that result from head motion.

A detailed account of the methods used to analyze the x,y fixation data from eye position recording was reported previously (13). For this study, if 50% of the area of a fixation cluster overlapped a target location (defined by an area of 0.5° radius surrounding the target) it was considered a "hit" (a true-positive response if the actual target was reported, a false-negative response if it was not). The same criterion was used for false-positive reports, except that the fixation cluster overlapped the erroneously reported nontarget location. True-negative decisions constituted those areas with fixation clusters that did not contain a target or a false-positive response (ie, scenic background).

RESULTS

Table 1 shows the AFROC A1 areas for finding NINA and WALDO. There was no statistical difference in the proportion of targets detected between radiologists and laypeople in either task. Consistent with this finding, AFROC analysis of overall detection performance in the NINA task resulted in A1 = .569 (standard deviation [SD] = .116) for radiologists and A1 = .689 (SD = .136) for laypeople. The difference was not significant (t_9 = 1.58). For the WALDO task, the estimated A1 for radiologists was .650 (SD = .086) and for laypeople was .690 (SD = .116). This difference was also not significant (t_{12} = 0.80).

Data from eye position recording were used to determine elapsed time until observers first fixated on NINA

Figure 1. Scanning pattern of lay person searching for the word "NINA" in a drawing called "The Apartment" by Al Hirschfeld. The lay person carried out a clockwise circumferential scan and fixated on the word NINA at 9 seconds. (Reprinted, with permission, from reference 10.)

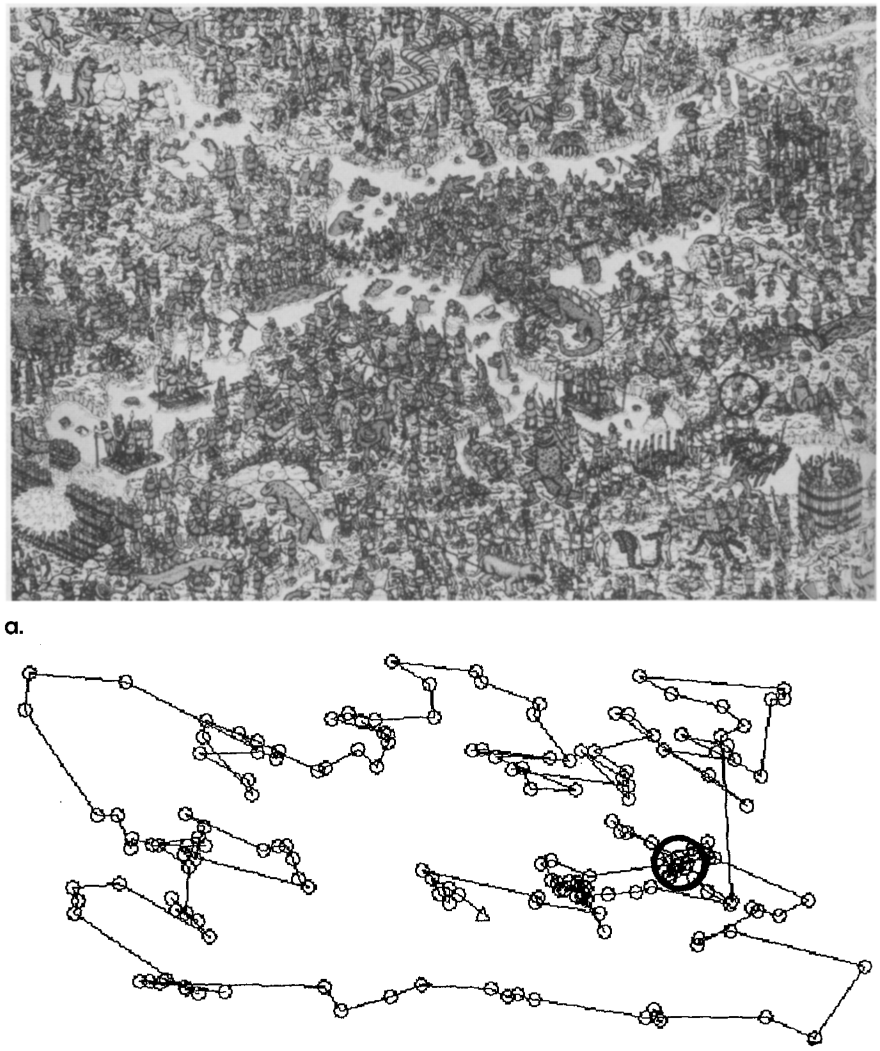
Figure 2. Scanning pattern of a radiologist searching the same scene as that use in Figure 1. The radiologist fixated on the spaghetti being strained by Jack Lemmon's tennis racket for 11.5 seconds before moving on. As a result, he did not fixate on the word "NINA" until 20 seconds into the search. Notice that even though the size of the NINA target is relatively large, because the letters are integrated into the structure of the lamp the target lacks peripheral conspicuity and therefore requires direct fixation to be detected.

or WALDO (a true-positive or false-negative response) after the search commenced (Table 2). Figure 1 shows the scanning pattern of a layperson, and Figure 2 shows the scanning pattern of a radiologist; both observers were searching for NINA in a scene taken from "The Apartment" by Al Hirschfeld (10). The radiologist's search pattern contains a greater density of fixations per scanning unit (ie, more detailed) and covers less of the image than the layperson's circumferential search pattern. This greater density was reflected in cumulative dwell time for various decision outcomes that was longer for radiologists than laypeople in all but one case.

This scanning strategy difference may account for the fact that laypeople were faster than radiologists in fixating on the NINA target ($F_{1,53} = 6.93$, $P < .01$). The difference in scanning strategy between radiologists and laypeople was not significant for WALDO. This may be because the experimenter gave the observers feedback about errors during the search so that they continued to search until they either found WALDO or time ran out. In the scene "Where's WALDO among the Monstrous Monsters" (Fig 3a) by Martin Handford (11), the layperson repeatedly fixated on WALDO (circled in the lower left corner) and reported finding it after a 23-second search. The radiologist carried out an extensive 2-minute search of the same scene, fixated on WALDO, but did not report finding him (Fig 4b). Figure 5 is a close-up of the WALDO scene.

The mean total viewing time was shorter for laypeople than radiologists in the WALDO task but not the NINA task (Table 3). Observers were given unlimited time of up to 2 minutes to search for WALDO. There were instances in both Hirschfeld and Handford test pictures where a target was not found. The Hirschfeld pictures contained multiple NINAs and Handford pictures contained only one WALDO; thus, rather than try to adjust arbitrarily the viewing times by adding a constant time to reflect misses, we simply eliminated the misses from the analysis. Regardless of whether an arbitrary time was added into the analysis, the mean total viewing time was significantly shorter for laypeople than radiologists only on WALDO pictures ($F_{1,110} = 5.46$, $P < .05$; arbitrary times for misses eliminated).

Table 4 shows the relationship between cumulative dwell time spent on a true or false target location and the correctness of the observer's decision about whether a true NINA or WALDO was or was not present at that location. Generally, observers in both NINA and WALDO tasks spent significantly more time dwelling on locations from which a positive decision was generated than on locations from which a true-negative decision was generated (Sheffe test, $P < .01$). In addition, when readers dwelled on locations from which a false-negative deci-



a.

b.

Figure 3. **(a)** Scene from *Where's WALDO among the Monstrous Monsters?* (11). The observers searched a full-color, 48 x 31-cm, poster-size drawing. The reduced black-and-white photographs give a false impression of the actual search task, but they do convey the density of pictorial detail present in the original. (Reprinted, with permission, from reference 11.) **(b)** The scanning pattern of a lay person. The layperson started the search near the lower middle of the picture (designated by the triangle) and reported finding WALDO after 23 seconds. WALDO is circled in the lower right corner of the picture and the scanning pattern. Note the density of fixations required to search the dense pictorial detail for WALDO.

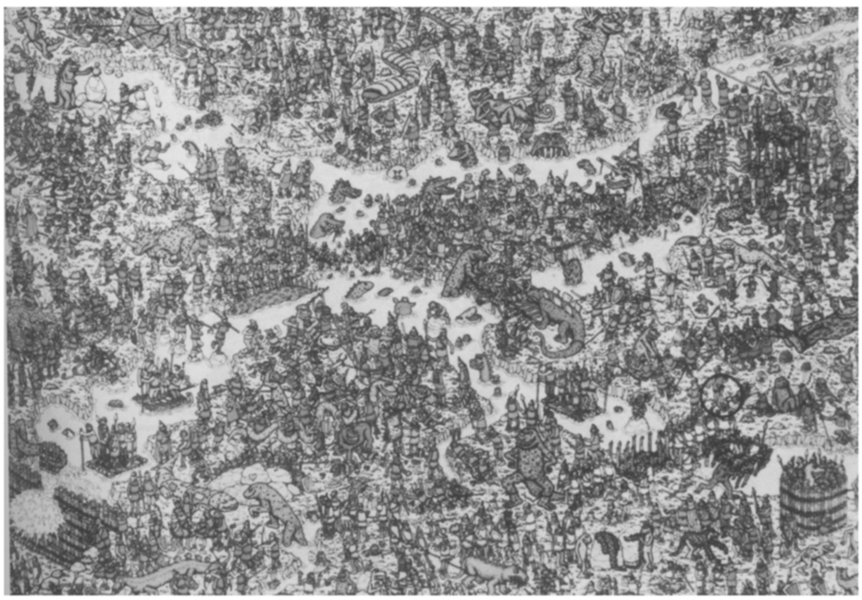
sion was generated, dwell time was significantly increased compared with dwell time on an area from which a true-negative decision was made (Sheffe test, $P < .01$).

DISCUSSION

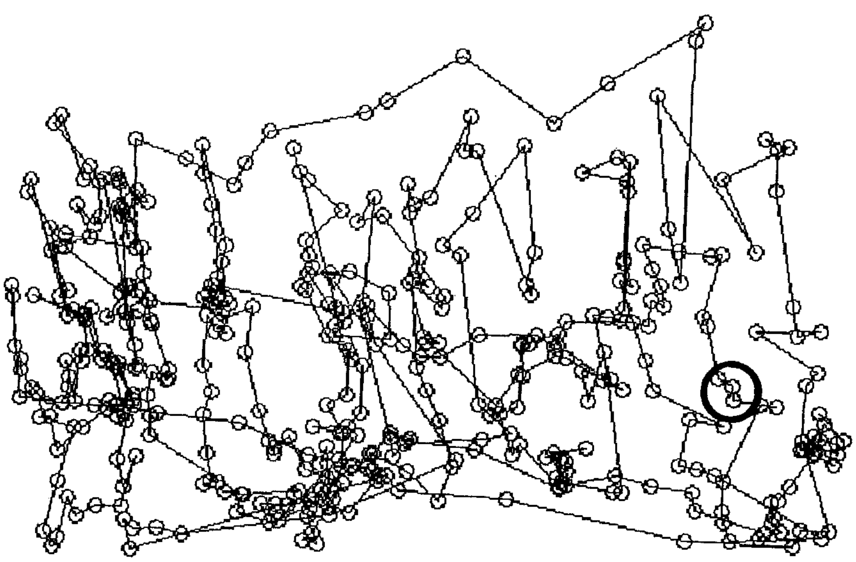
There have been a number of attempts to try to correlate the diagnostic ability of radiologists with a variety of perceptual tasks (1–5). Some have been successful and some have not. Few, if any, studies have compared the performance of radiologists to that of laypeople on visual tasks that emulate what the radiologist does while searching and interpreting a medical image for an abnormality. We used two art search tasks that we believe captured many of the characteristics of radiologic search but did not require special training or experience to perform. If radiologists were better searchers and analyzers than laypeople, either innately or by specific training, the hypothesis was that the radiologists would perform better at the generalized search task. In fact, we discovered that the radiologists were no better at the general search task than laypeople. What does this mean?

First, we assumed that the art search tasks tap similar basic perceptual and cognitive skills of visual search, detection, and interpretation that radiology tasks do in searching for abnormalities. This may not be the case, but before we accept this conclusion let us look at a second possibility.

Second, this study can be viewed as expanding on the nature of radiology expertise and how it transfers from one task to another. Let us assume that the art image search task tapped perceptual and cognitive skills similar to those



a.



b.

Figure 4. **(a)** The same scene as that shown in Figure 3a. (Reprinted, with permission, from reference 11.) **(b)** The scanning pattern of a radiologist. The radiologist began the search in approximately the same location as the layperson but did not find WALDO (false-negative result) during the 2-minute search period, even though he did fixate on WALDO (as indicated by the circle) toward the end of the search.

of the radiologic image search task, but that both radiologists and laypeople lacked experience searching and interpreting art targets. This would have led to the same pattern of results as our first conclusion. We know from previous research that radiology expertise depends heavily on the interaction of experience with training. For example, Nodine et al (6) showed that it takes a 13- to 200-fold increase in experience to effectively improve mammography screening performance during mammography training. Beam et al (14) suggested that this range of experience may be underestimated by at least 10-fold and that because of the relatively low occurrence of breast cancer, radiology residents rarely encounter a case of breast cancer during a clinical mammography rotation (14).

We know from a number of studies that radiologists search x-ray images more effectively than do nonradiologists. For example, reanalysis of Kundel and La Follette's 1972 study (8) shows that significantly fewer fixations were required to detect and correctly report lung lesions by radiologists and radiology residents than medical students (mean = 5.21 fixations for radiologists and residents vs 13.27 fixations for medical students, $F_{1,23} = 5.76$, $P < .05$). In this case, search efficiency was reflected by length of the scanning pattern required to

Figure 5. Detail of part of the scene from *Where's WALDO among the Monstrous Monsters?* (Reprinted, with permission, from reference 11.)

sample and report the lesions correctly. This pattern of results has been repeatedly replicated (6,8,15,16).

We hypothesize that because radiologists lacked experience searching for art targets, their radiology expertise did not positively transfer for the limited art-testing experience. This finding confirms what the well-known learning theory of Osgood predicted long ago—namely, that degree of transfer depends on the similarity of training and test situations (17). The similarities in task requirements may have been outweighed by the manner in which perceptual and cognitive processes interact in finding and disembedding novel target features from art image backgrounds compared with x-ray image backgrounds. What is critical in transfer from radiology to art tasks is the observer's understanding about how the pictorial background acts to camouflage the target, and this understanding requires a great deal of experience detecting, recognizing, and deciding that a target has been found. For the radiology task of searching for a lesion in a chest or breast x-ray image, lesions are camouflaged primarily by occlusion and blending of the lesion with anatomic background structures like blood vessels on end or dense breast parenchyma. In the case of searching Hirschfeld's drawings, NINA is camouflaged by blending the letters of the name into background scenery that contains features designed to mimic alphabet letters. In the case of Handford's drawings, WALDO is camouflaged primarily by mimicry. Subtle variations in the color patterns and shapes that are distinctively assigned to WALDO are also used to create foils. Thus, because different tasks call on different perceptual mechanisms for detecting and recognizing targets, what we may have observed in the present study is a low degree of perceptual learning transfer by the radiologists so that they performed at the level of inexperienced laypeople. In fact, our data show that radiologists tended to find fewer art targets and miscalled more art targets falsely than laypeople. From the standpoint of transfer of radiology expertise, neither perceptual discrimination nor visual search skills carried over to the art tasks.

Finally, analysis of eye position data revealed that when both radiologists and laypeople missed art targets, they typically spent significantly more visual dwell time fixating on the true target than negative, nontarget background locations on the images. This finding, together with the ranking of dwell times associated with true- and false-positive decisions, has also been observed in visual search tasks in radiology (7,18,19). Thus, it seems that at least in this respect, the art image task was tapping fundamental perceptual processes associated with visual search, detection, and decision making.

These data have a couple of important implications for testing and training. The first implication follows from our conclusion about transfer: Radiologists may not be superior visual searchers and analyzers in a general sense. They may be expert at searching radiologic images (8), but their search and analysis skills do not transfer to new tasks that have similar requirements. If this is true, then

Table 3
Total Viewing Time Spent in Search for NINA or WALDO Targets

	NINA		WALDO	
Viewing Time	Radiologists	Laypeople	Radiologists	Laypeople
Mean	44.90 (11)	44.66 (15)	61.42 (55)	48.02 (57)
SD	21.48	17.11	32.13	28.48

Note.—Search time data are given in seconds. Numbers in parentheses are number of records. SD = standard deviation.

Table 4
Cumulative Dwell Time Associated with Various Decision Outcomes for NINA and WALDO Test Pictures

Decision Outcome	NINA		WALDO	
	Radiologists	Laypeople	Radiologists	Laypeople
True-positive				
Mean	2,525 (14)	1,393 (17)	1,775 (55)	1,225 (57)
SD	1,315	981	1,354	676
False-negative				
Mean	1,340 (8)	1,223 (7)	2,773 (15)	2,046 (13)
SD	911	825	1,425	1,214
False-positive				
Mean	ND	ND	1,585 (26)	1,475 (20)
SD	ND	ND	800	749
True-negative				
Mean	798 (64)	521 (57)	937 (9,736)	993 (10,421)
SD	806	599	1,641	1,475

Note.—Dwell time data are given in milliseconds. Numbers in parentheses are number of responses. ND = no data, SD = standard deviation.

this finding has direct consequences on a second implication, training—that is, selection of residents for radiologic training (and developing tests for this selection process) and methods of training during radiology residencies. Freundlich and Murphy (20) found that 93.5% of medical students taking a radiology elective expected to be able to correlate their interpretations of radiographs and other medical images with radiographic reports. However, did they consider what happens when a disagreement occurs? Obviously not. Even more surprising was the finding that many medical students believed that a 4-week elective adequately prepared them to interpret radiographs independently. In all probability radiology residency programs do not share this view. In fact, there are efforts being made to change the radiology residency curriculum to better prepare residents for a career in radiology (21–23). The main question, of course, is exactly what and how do we teach residents to be expert radiologists? Our results suggest that perceptual skills for radiology require knowing what distinctive features differentiate abnormal from normal anatomic structures (through medical training) and knowing how these features are transformed by radiographic imaging and interpreted within the context of diagnostic hypothesis testing and problem solving (through radiologic experience). As our results suggest, these skills may in fact be specific to the situation of interpreting x-ray images and may not generalize to other nonradiologic hide-and-seek search tasks.

The testing and training issue is also interesting in light of the fact that many training institutions may have to decrease the number of radiology residents in the near future (24). Our study raises questions about the effectiveness of testing programs to predict which medical students would make good radiologists. Our findings show how difficult it is to develop a testing situation to predict how much perceptual learning carries over from radiology search to visual tests. On a generalized search and analysis task, radiologists are no better than laypeople. Bass and Chiles (1) found that performance on perceptual tests had little correlation with diagnostic accuracy in detecting pulmonary nodules. One group (2–5) did find a good correlation between perceptual test performance and ratings of residents' diagnostic skills, but one of the studies (3) found that the correlation was poor during the 1st year and stronger after that.

These studies differ from the present study in that they did not compare the performance of radiologists with that of laypeople. They looked only at the performance of radiologists and those training to be radiologists. Therefore, the effects of training may have already influenced their skills to some degree. It is impossible to tell whether the observers tested had different or better perceptual skills on entering their residency, or whether the training enhanced or fostered already existing perceptual skills that had not previously been tapped. Our study tested the performance of radiologists and laypeople on a visual search task and found little difference in performance. This suggests that if radiologists do possess superior search skills, they may be specific to the radiologic search task and may not be evident on other types of search tasks that do not deal with x-ray images.

REFERENCES

1. Bass JC, Chiles C. Visual skill: correlation with detection of solitary pulmonary nodules. Invest Radiol 1990; 25:994-998.
2. Smoker WRK, Berbaum KS, Luebke NH, Jacoby CG. Spatial perception testing in diagnostic radiology. AJR 1984; 143:1105-1109.
3. Berbaum KS, Smoker WRK, Smith WL. Measurement and prediction of diagnostic performance during radiology training. AJR 1985; 145:1305-1311.
4. Berbaum KS, Platz C. Perception testing in surgical pathology. Hum Pathol 1988; 19:1127-1131.
5. Smith WL, Berbaum KS. Improving resident selection: discrimination by perceptual abilities. Invest Radiol 1991; 26:910-912.
6. Nodine CF, Kundel HL, Lauver SC, Toto LC. Nature of expertise in searching mammograms for breast masses. Acad Radiol 1996; 3:1000-1006.
7. Krupinski EA. Visual scanning patterns of radiologists searching mammograms. Acad Radiol 1996; 3:137-144.
8. Kundel HL, LaFollette PS. Visual search patterns and experience with radiological images. Radiology 1972; 103:523-528.
9. Nodine CF, Carmody DP, Kundel HL. Searching for NINA. In: Senders JW, Fisher DF, Monty RA, eds. Eye movements and the higher psychological functions. Hillsdale, NJ: Lawrence Erbaum, 1978; 241-258.
10. Hirschfeld A. The world of Hirschfeld. New York, NY: Abrams, 1970.
11. Handford M. Where's WALDO? the magnificent poster book. New York, NY: Little, Brown, 1991.
12. Chakraborty D, Winter LHL. Free-response methodology: alternative analysis and a new observer-performance experiment. Radiology 1990; 174:873-881.
13. Nodine CF, Kundel HL, Toto LC, Krupinski EA. Recording and analyzing eye-position data using a microcomputer workstation. Behav Res Meth Instrum Comput 1992; 24:475-485.
14. Beam CA, Layde PM, Sullivan DC. Variability in the interpretation of screening mammograms by US radiologists. Arch Intern Med 1996; 156:209-213.
15. Gale AG, Walker GE, Roebuck EJ, Wilson ARM, Worthington BS. Mammographic screening: what do radiologists look for? (abstr). Br J Radiol 1990; 63:S52.
16. Lesgold AM, Rubinson H, Feltovitch P, Glasser R, Klopfer D, Wang Y. Expertise in a complex skill: diagnosing x-ray pictures. In: Chi MTH, Glaser R, Farr MJ, eds. The nature of expertise. Hillsdale, NJ: Lawrence Erbaum, 1988; 311-342.
17. Osgood CE. Method and theory in experimental psychology. New York, NY: Oxford University Press, 1956; 532.
18. Kundel HL, Nodine CF, Carmody DP. Visual scanning, pattern recognition and decision-making in pulmonary nodule detection. Invest Radiol 1978; 13:175-181.
19. Kundel HL, Nodine CF, Krupinski EA. Computer-displayed eye position as a visual aid to pulmonary nodule interpretation. Invest Radiol 1990; 25:890-896.
20. Freundlich IM, Murphy WA. Medical students who choose a radiology elective: career decisions, motivations, and intentions. Acad Radiol 1995; 2:527-532.
21. Mundy WM, Binet E. A comprehensive objective-based curriculum for radiology residents. Acad Radiol 1995; 2:173-178.
22. Ekelund L, Lanphear J. Diagnostic radiology in an integrated curriculum: experience from the United Arab Emirates. Acad Radiol 1997; 4:653-656.
23. Rao VM. A perspective on radiology residency curriculum guidelines: results of the 1995 survey of program directors. Acad Radiol 1996; 3:512-516.
24. Manaster BJ. Decreasing the number of radiology residents: the impact on radiology departments and resident education. Acad Radiol 1995; 2:1113-1114.

Announcement

The University of Chicago will host the **First International Workshop on Computer-Aided Diagnosis (CAD)** on September 20–23, 1998, at the University of Chicago Downtown Center, Chicago, Illinois. The meeting will provide a forum for the leading researchers and practitioners in CAD and will encompass automated image analysis, quantitation of image information, two- and three-dimensional multimodality image integration, advanced image processing, and artificial neural network applications. Sessions will include new developments in chest, breast, vascular, and three-dimensional/CT/multimodality imaging. Related developments in digital image acquisition and picture archiving and communication systems, or PACS, will also be addressed. The sponsoring chairman is Martin J. Lipton, MD, and the organizing committee consists of Kunio Doi, PhD, Heber MacMahon, MD, Maryellen L. Giger, PhD, and Kenneth R. Hoffmann, PhD. Attendance will be limited to 150 attendees on a first-come basis. The registration fee is $400.

For more information, contact the International Workshop on Computer-Aided Diagnosis, The University of Chicago, Department of Radiology MC2026, 5841 S Maryland Ave, Chicago, IL 60637; e-mail: *cad@uchicago.edu;* fax: 773-702-0371.