

HOW DO RADIOLOGISTS USE THE HUMAN SEARCH ENGINE?

Jeremy M. Wolfe^{1,2,*}, Karla K. Evans³, Trafton Drew⁴, Avigael Aizenman¹ and Emilie Josephs⁵

¹Visual Attention Laboratory, Brigham and Women's Hospital, Cambridge, MA, USA

²Departments of Radiology and Ophthalmology, Harvard Medical School, Boston, MA, USA

³Department of Psychology, University of York, York, UK

⁴Department of Psychology, University of Utah, Salt Lake City, UT, USA

⁵Department of Psychology, Harvard University, Cambridge, MA, USA

*Corresponding author: jwolfe@partners.org

Radiologists perform many 'visual search tasks' in which they look for one or more instances of one or more types of target item in a medical image (e.g. cancer screening). To understand and improve how radiologists do such tasks, it must be understood how the human 'search engine' works. This article briefly reviews some of the relevant work into this aspect of medical image perception. Questions include how attention and the eyes are guided in radiologic search? How is global (image-wide) information used in search? How might properties of human vision and human cognition lead to errors in radiologic search?

INTRODUCTION

Many of the most important tasks performed by radiologists can be characterised as visual search tasks in which the radiologist looks for one or more instances of one or more types of target item in a complex scene containing various distractor items. Cancer screening tasks are obvious examples of such radiological search tasks. To understand how radiologists do these tasks, it must be understood how the human 'search engine' works, more generally. Just as your internet browser has a search engine that allows you to specify the current object of search, your visual system comes with a search engine that tries to guide your attention to the current object of your visual attention. Obviously, the rules governing these search mechanisms are rather different from each other. The present authors' research is devoted to uncovering the principles of the human search engine and their application by experts like radiologists.

Not all medical image perception tasks have a prominent search component. If the radiologist is asked to interpret the meaning of the contents of some specific location in an image, performance will be constrained by human visual capabilities but not by the properties of the human search engine. Thus, if the radiologist is asked if he sees a fracture in this location or a tumour in that location, the answer will depend on acuity and contrast sensitivity⁽¹⁾. If, however, the radiologist is asked to examine an abdominal computed tomography (CT) examination to determine if this crash victim suffered any internal injuries or if there are any metastases in the liver, then the answer will depend not only on vision but also on the ability to direct attention to the correct locations in the image. Only then can the interpretive processes be put to work.

This article will review several aspects of human visual search performance as they apply to radiology. In some cases, the data come from the testing of radiologists. In other cases, the questions are raised in the context of radiology, but the data are collected from non-expert observers who are more available for the sort of extensive testing that may be required. The application of data from non-experts to the work of radiologists is based on the conviction that radiologists are experts who have learned to use the human search engine to perform specialised searches. Nevertheless, they remain subject to the rules and limitations of that search engine.

A series of questions will be addressed:

- (1) How are attention and the eyes guided when radiologists search?
- (2) How might this guidance lead to errors in search?
- (3) What non-selective information can be used by radiologists?
- (4) How does the low prevalence of targets in screening tasks influence performance on those tasks?

HOW ARE ATTENTION AND THE EYES GUIDED IN VISUAL SEARCH?

Guided search

Search is necessary because humans cannot fully process everything in the visual field at one time. Object recognition is limited to one or, perhaps, to a small number of objects at one time⁽²⁾. Attention is not deployed randomly. Instead, it is guided to the likely target locations. Guidance can be broadly divided into three varieties. Attention is attracted to items and/or locations in the visual field that are 'salient' in a bottom-up manner. 'Bottom-up' salience

refers to properties of the stimulus that do not depend on the observer's intentions. It is based on local differences in basic features^(3, 4, 5). Thus, for example, a red item among green or a moving item among stationary items will 'pop out' of a display^(6, 7) in a bottom-up manner. It does not matter if the observer is particularly interested in that red item. It will tend to attract attention, regardless. See Wolfe and Horowitz⁽⁸⁾ for a discussion of the limited set of attributes that pop out.

These attributes are sometimes categorised as 'pre-attentive' since they appear to be processed across the entire visual field without the need to use attention to select individual items. Some other properties of the stimuli, while they may not produce pop-out by themselves, can guide attention by a more indirect route. For example, manipulating apparent depth can alter the apparent size of items and an item that now appears larger will attract attention, because size is one of the basic attributes that can guide attention⁽⁹⁾. A particularly interesting modulator is 'value'. A property will come to attract attention more vigorously if it is made valuable (e.g. by association with a monetary reward). Thus, if you pay people more to find red things than green things, they will learn to guide attention to items of the more valuable colour^(10, 11).

The second variety of attentional guidance is top-down, user-driven guidance. 'Top-down' refers to the state of the observer. When a searcher has some representation of a target or targets in mind and guides attention to those items that share basic features of that target, that is top-down guidance. Thus, if someone is looking for a red letter, 'R', she will guide attention to red letters and ignore letters of other colour. For a radiologist looking for lung nodules in a chest CT, attention will be guided in a top-down manner towards items that are small and white and round. It is important to note that the representation of the target serves two roles. One role is the top-down guiding role. The second role is a decision role. Once attention is attracted to an item, a decision must be made as to whether the selected item matches the target representation. The guidance is limited to the basic features of the target. Is the item small, white and round? These guiding features are quite crude and categorical⁽¹²⁾. At the decision stage, it can be determined if this item is a possibly cancerous lung nodule based on much more subtle matching of visual stimulus to remembered template.

Bottom-up and top-down guidance are based on the properties of individual items. The third form of guidance is scene guidance; guidance based on an understanding of the scene contents and layout. For example, if you are looking for people in Figure 1, you will guide your attention to locations where people are most likely to be⁽¹³⁾. You know that people do not float in the air and require a roughly horizontal surface to support them. These physical constraints are called 'syntactic guidance'^(14, 15). Moreover, based on



Figure 1. Syntactic and semantic guidance help to find likely locations for people in real-world scenes like this.

the vast experience exploring the world, it is known that a human is more likely to appear in the scene in Figure 1 than a giraffe. This is called 'semantic guidance', guidance by the meaning of the scene.

Scene guidance is learned, and this can be seen in changes in the eye movements of radiologists as they increase in their expertise^(16, 17). Characteristically, trainee radiologists make more eye movements when evaluating an image than do experts, and those eye movements cover more of the area of the image. One might think that it would be desirable to look everywhere, but it is no more desirable to spend time looking for lung nodules where lung nodules do not occur than it would be to spend time looking for giraffes in Gothenburg, Sweden (the location of the scene in Figure 1). The reduction in the number and distribution of eye movements in expert radiologists probably reflects a combination of all of these forms of guidance. By developing a better template for the target of search, areas of the image can be removed from consideration. If you are looking for a dark vertical object in Figure 2, there is no point deploying attention or the eyes to light or tilted items. If a radiologist is looking for lung nodules, there are some portions of the image that will not be examined because they do not typically contain lung nodules and there will be other regions that will not be examined because they lack the preattentive features that could turn out to be nodules if scrutinised.

Moreover, one can learn to ignore irrelevant, bottom-up signals. While attention is attracted to salient features in a scene, if these are known to be irrelevant, the tendency to attend to those signals can be suppressed, at least to some extent. Were this not the case, it would be found that attention would be continually pulled to salient irrelevancies like bright highlights on shiny surfaces. It is not possible to completely suppress the attention to salient stimuli. Some types of stimuli, like the onset of new objects, seem to be particularly potent in their ability to capture

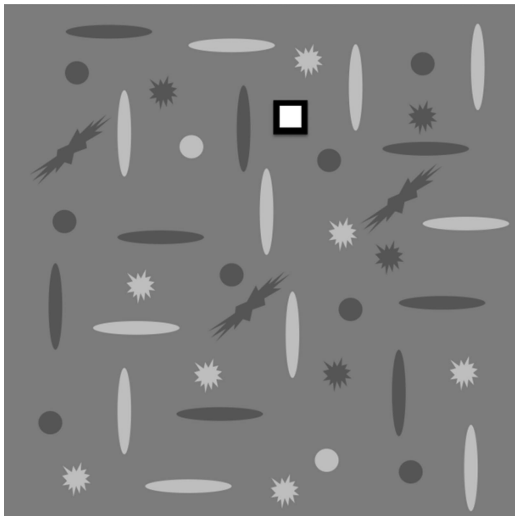


Figure 2. Illustration of bottom-up and top-down guidance. One item ‘pops out’ because of its bottom-up saliency. Gray, round ‘nodules’ can be found by guiding attention in a top-down manner to gray and round items.

attention⁽¹⁸⁾. If given enough practice, no type of capture appears to be absolute⁽¹⁹⁾. However, it seems likely that, if liberated from the limits of what can be done to people in laboratory settings, there will always be some stimulus that will capture attention in a bottom-up manner or, as Sully⁽²⁰⁾ put it in the 19th century, ‘One would like to know the fortunate (or unfortunate) man who could receive a box on the ear and not attend to it’ (p. 146 of Ref. ⁽²⁰⁾).

These ideas can be summarised in the ‘Guided Search’ model^(21, 22, 23). Guided Search would propose that the radiologist has in mind a target; say, lung nodules. This representation gets translated into top-down guidance towards small, white and round items. Bottom-up attraction to the irrelevant red light-emitting diode (LED) on the monitor would be suppressed, at least somewhat. The expert’s knowledge provides scene guidance to relevant parts of the image. All these sources of information are combined to create a ‘priority map’⁽²⁴⁾ that represents the observer’s best guess as to the likely location of targets. Attention is then directed to items in the order of their current priority. Items are selected, one after the other, and matched against the internal representation of the target so that the observer can make a decision: is this attended item a lung nodule or not?

There are many aspects of Guided Search that are beyond the scope of the present review. For example, how fast are items selected? How long does it take to process a single item⁽²⁵⁾? How does the priority map evolve over the course of a search? It must change as the eyes move about the image since the saliency of items

will change with their distance from the present fixation^(26, 27). How can observers avoid returning to salient locations that do not contain targets. It was originally proposed that each rejected location was inhibited^(21, 28), but this form of extensive ‘inhibition of return’ does not seem to operate in search⁽²⁹⁾ and more nuanced accounts of how the system avoids perseveration have been developed⁽³⁰⁾.

One last topic that will only briefly be mentioned here is the possibility of searching for multiple targets at the same time. For instance, that a radiologist looking for small white round nodules is also likely to be looking for a list of other possible findings. Search of a visual stimulus for multiple targets, held in memory, is known as hybrid search⁽³¹⁾. It turns out that humans are capable of searching for many types of target at the same time, though they cannot guide their attention as effectively to all the features of multiple target types. These and other topics are reviewed more extensively elsewhere^(32, 33, 34, 35).

WHEN MIGHT GUIDANCE PRODUCE ERRORS?

The processes of guidance, vital as they are, sometimes cause problems. The present authors conducted a study of the eye movements of radiologists as they searched for lung nodules in a 3D stack of CT images. On the last case, a gorilla was inserted into the lung (Figure 3). Fully 20 of 24 radiologists failed to notice this⁽³⁶⁾. A gorilla, of course, is not a lung nodule, but it is the sort of ‘incidental finding’ that one might hope would attract attention. Why might it be missed? It is not because the radiologists were in any way negligent or inattentive to the task at hand. They were using their human search engine, and that search engine is subject to ‘inattention blindness’⁽³⁷⁾. The authors’ use of a gorilla was a homage to the classic inattention blindness demonstration of Simons and Chabris⁽³⁸⁾. In that experiment, observers were watching a basketball game and were instructed to count the number of times that the white team passed the ball. An actor in a gorilla suit wandered into the middle of this game, waved and left. About 50 % of viewers did not report the gorilla, even when asked if they had seen anything unusual. This chest CT experiment simply translated this experiment into the radiology domain. It can be seen as the downside of guidance. The radiologists were searching for small white nodules. In effect, they guided attention away from big black gorillas and worked to suppress any bottom-up saliency signals produced by the gorilla. Thanks to top-down guidance, radiologists would have done better if the gorilla had been white. The better the match between the features of an odd item and the features of the target, the less severe the inattention blindness⁽³⁹⁾. Radiologists can be blind to more naturalistic stimuli as Potchen *et al.*⁽⁴⁰⁾ found

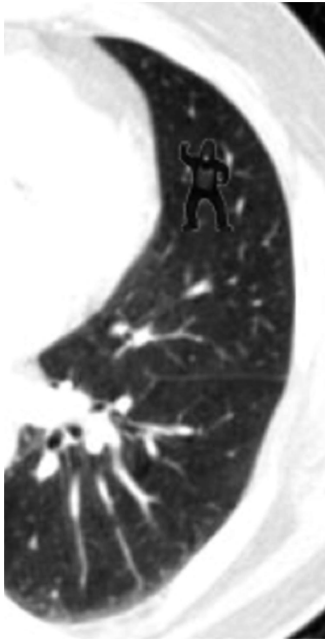


Figure 3. An unusual piece of a lung CT.

when they removed the clavicle from an image, only to have its absence missed. That particular example also serves to illustrate the difficulty in detecting the absence of a stimulus compared with detecting its presence⁽⁴¹⁾. For a similar example, see Lum *et al.*⁽⁴²⁾.

The ‘inattentive’ aspect of inattention blindness can be seen in the eye movements of the radiologists in the study. In some cases, it might be argued that the gorilla was missed because it was never fixated so its image never landed on the fovea. In many cases, however, this was not the explanation. On average, the gorilla was fixated for about a second over the course of a trial; more than sufficient to decide whether an object is a target⁽⁴³⁾. Thus, while most of the radiologists looked right at the gorilla, it was not consciously perceived due to the inherent limitations of the human search engine.

How are attention and the eyes guided in radiologists?

Beyond showing that radiologists are not immune to inattention blindness, the authors’ primary interest in the experiment described above was to look at the eye movements of radiologists as they searched through the 3D volume of image data that constituted the lung CT. There is a small, influential literature on eye movements in the XY plane as radiologists examine classic 2D images^(44, 45, 46, 47, 48). What happens when radiologists can also control where they are fixating in the z-direction by scrolling up and down through a

stack of CT ‘slices’ through the chest? It was found that the behaviour fell into two broad categories, which were dubbed ‘drillers’ and ‘scanners’. Drillers kept their eyes largely in one quadrant while moving through the depth of the stack. Then, they would move to a different XY location and drill back up through the stack, and so on. Scanners tended to look at many positions in the XY plane while holding depth constant. Then, they would move a small distance in depth and scan the XY image again⁽⁴⁹⁾.

Anecdotally, radiologists describe learning one or the other technique from their mentors, but is one method superior? While it was found that nodule detection was significantly higher for drillers, the sample ($N=24$) was too small to be sure. Accordingly, an analogue of the task was developed for non-experts. Observers scrolled through a 3D volume of noise while looking for ambiguous letters. Ts were targets. Ls were not. Observers were instructed to be drillers or scanners. While the data showed that they followed the instructions, no difference between behavioural performance could be seen⁽⁵⁰⁾. The result has the problem of a negative result. It could be that a different version of the experiment would give the advantage to one or the other method. For the present, however, it seems that both methods serve to encourage some form of systematic examination of the display and that either can be effective.

NON-SELECTIVE INFORMATION IN SEARCH

When talking to radiologists about screening tasks like mammography, experts will sometimes report that they get a feeling about a case from the first glimpse. They seem to sense that something is present before they find it. Forty years ago, Kundel⁽⁵¹⁾ spoke of ‘Interpreting chest radiographs without visual search’ and incorporated the idea of global or ‘gestalt’ processing in radiographic interpretation⁽⁵²⁾. Like guidance and inattention blindness, this ability to extract meaning from a complex scene in a brief glimpse is an aspect of normal visual capabilities that radiologists have fine-tuned and applied to more specialised searches, through their experience with radiographs. Humans are able to extract the ‘gist’ of a scene in a fraction of a second. Thus, in <100 ms, observers can determine if a scene is natural or man-made or if a scene is navigable or not⁽⁵³⁾.

The present authors conducted a study in which it was asked if mammographers could look at a bilateral mammogram for less than a second and determine if the woman should be called back at above chance levels. Subtle cancers were deliberately used: masses and architectural distortions, rather than bright white calcifications. Nevertheless, the experts could reliably make this decision at above chance levels when given a quarter of a second exposure to the image [$d' = \sim 1.0$, area under the curve (AUC) = ~ 0.75]⁽⁵⁴⁾. The

observers were also presented with an outline of the breasts and asked to mark their best guess for the location of a lesion, if one were present. In this task, the experts performed at chance. That is, there were times when they correctly detected that something was there, but they had no idea where.

PREVALENCE EFFECTS

How much scrutiny does the current scene deserve? This question needs to be answered constantly throughout the day: How long should one look for the perfect tomato in this display at the farmers' market? At what point should one give up looking for the keys in the purse and look somewhere else? Radiologists must develop an answer specific to their domain of expertise. How long should one look for pathology in this lung? There is a lot to look at in a 300-slice CT, but there are more patients waiting in the queue. Screening tasks such as screening for breast or lung cancer represent an important class of radiological search that comes with a distinctive feature: the targets are very rare. For instance, in mammography, the prevalence of disease in a North American screening population is on the order of 3 or 4 per thousand women screened⁽⁵⁵⁾. Target prevalence in laboratory studies of search is typically 50 % (often more). Does prevalence matter? It has long been known from studies of vigilance that rare events are missed more often than common ones⁽⁵⁶⁾, but in vigilance experiments, observers are typically asked to look for something, like a flash of light that is presented and then vanishes. A mammogram does not vanish. It is present until the radiologist chooses to dismiss it. What is the impact of prevalence in this case?

Low prevalence is also a characteristic of many non-medical screening tasks, like airport baggage screening. Indeed, in the present authors' first work in this area, baggage X-ray stimuli were used. As shown at the top of Figure 4, with non-expert observers searching for guns and knives, it was found that false-negative (miss) error rates were about twice as high at 2 % prevalence as they were at 50 %⁽⁵⁷⁾. With the complex, possibly ambiguous stimuli, the primary cause of this effect was a criterion shift⁽⁵⁹⁾. When targets were rare, observers were more likely to reject an ambiguous target. When targets were very common (e.g. 98 % prevalence in Wolfe and Van Wert⁽⁵⁹⁾), observers became much more likely to label an ambiguous item as a target. Importantly, the same items could be missed in low and found in high-prevalence situations. It is not the appearance of the stimulus that is changing. Rather, it seems that it is the unconscious decision rule that is changing. Observers also become faster to declare themselves to be finished with an image under low-prevalence conditions, but simply forcing them to slow down does not make

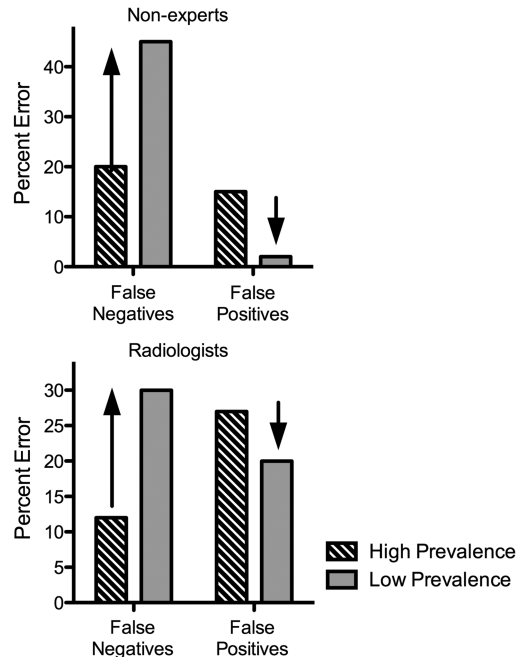


Figure 4. Effect of target prevalence on error rates in non-expert observers (top panel) and in radiologists performing a screening mammography task (bottom panel). Data replotted from Wolfe *et al.*⁽⁵⁷⁾ and Evans *et al.*⁽⁵⁸⁾.

observers markedly less likely to reject low-prevalence targets⁽⁶⁰⁾.

Thus, with non-expert observers, low prevalence increases miss errors and causes observers to terminate searches more quickly, but does this apply to experts, or does expertise immunise observers against these effects? The experiments that the authors ran with non-experts involved observers looking at thousands of images because, at low prevalence, you need thousands of images to get a reasonable number of target-present cases. This is not practical with radiologists. Instead, 100 cases were taken, 50 with and 50 without fairly subtle cancer, and slipped into regular practice at a large urban hospital (with appropriate consent from the radiologists in the women's imaging service). Less than one case per day was added so that not much of a burden was imposed on the practice, and the baseline prevalence was not changed very much. For comparison with that low-prevalence arm of the study, 6 radiologists were asked to read the 100 cases in a single block of 50 % prevalence.

As shown in the lower panel of Figure 4, the error results mirrored those that were obtained with non-experts. The radiologists had a false-negative rate of 12 % for these cases at high prevalence. At low prevalence, the false-negative rate was 30 %. That difference in percentage of cancers that were missed was highly

significant ($\chi^2_{(1)} = 11.77, p < 0.005$)⁽⁵⁸⁾. The false-positive errors showed the opposite pattern with 20 % false positives at low prevalence and 27 % at high prevalence, though that difference was not statistically significant ($\chi^2_{(1)} = 3.04, p > 0.05$). Similar results were found with cytologists reading cervical cancer screening slides⁽⁶¹⁾ and with newly trained transportation security officers, looking at luggage X-rays⁽⁶²⁾. In all cases, rare targets were missed more often.

Even though they had given consent, radiologists did not know *when* they would see a case in the low-prevalence arm of the study. Those cases simply appeared as part of the regular workflow. Thus, it is interesting that they missed more cancers while doing their ‘real’ job than they missed when participating in the study reading cases in the high-prevalence condition. Indeed, the same case that was missed in the low-prevalence situation would be found in the high-prevalence situation. These results should not be seen as casting any aspersions on the professionalism or expertise of the radiologists in the present study. Humans are built to be conservative about positive evidence for rare events. When this makes you less likely to assert that you have seen aliens landing in the backyard, that conservatism is probably a good thing. When it causes an expert to miss a rare but important target, then it becomes a problem worth studying and worth trying to correct.

THE CHALLENGE OF MEDICAL IMAGE PERCEPTION

The examples of radiologic search, described here, make it clear that there are a host of interesting questions raised by the interaction of the human search engine with the tasks of clinical radiology. Is drilling actually superior to scanning? Is that missed gorilla related to ‘satisfaction of search’ errors in radiology practice^(63, 64)? Is the non-selective signal, present in mammograms, something that could be used to improve performance in mammography? Can we counteract the criterion shift at the heart of the prevalence effect? All of these are good research questions, but they are difficult to study. Radiologists are a limited resource as experimental observers. They simply do not have the time to participate in experimental settings, and there is a limit to how one can manipulate clinical practice in the service of research. It is possible to address many of these questions using non-expert observers. However, even though experts and non-experts are working with the same basic search engine, conclusions based on studies of naïve observers are unlikely to be accepted by clinicians without corroborating results from radiologists. Indeed, there is some reluctance in journal and grant reviewing to believe that there is any value in addressing problems in medical image perception by testing non-experts. It is worth working on these structural

problems because, as long as humans are a vital part of medical image analysis, understanding the capabilities and limits of those humans will be central to improving outcomes in radiology.

FUNDING

This work was supported by the National Institute of Health (EY017001 to J.M.W.), the Office of Naval Research (N000141010278 to J.M.W.) and the National Science Foundation (SBE-0354378 to J.M.W.).

REFERENCES

1. Burgess, A. E. *Spatial vision research without noise*. In: The Handbook of Medical Image Perception and Techniques. Samei, E. and Krupinski, E. A., Eds. Cambridge University Press (2010).
2. Bricolo, E., Gianesini, T., Fanini, A., Bundesen, C. and Chelazzi, L. *Serial attention mechanisms in visual search: a direct behavioral demonstration*. *J. Cogn. Neurosci.* **14**, 980–993 (2002).
3. Braun, J. *Visual search among items of different salience: removal of visual attention mimics a lesion of extrastriate area V4*. *J. Neurosci.* **14**, 554–567 (1994).
4. Nothdurft, H. C. *Texture segmentation and pop-out from orientation contrast*. *Vision Res.* **31**, 1073–1078 (1991).
5. Nothdurft, H. C. *Feature analysis and the role of similarity in preattentive vision*. *Percept. Psychophys.* **52**, 355–375 (1992).
6. Egeth, H., Jonides, J. and Wall, S. *Parallel processing of multielement displays*. *Cognitive Psychol.* **3**, 674–698 (1972).
7. Treisman, A. *Features and objects in visual processing*. *Sci. Am.* **255**, 114–125 (1986).
8. Wolfe, J. M. and Horowitz, T. S. *What attributes guide the deployment of visual attention and how do they do it?* *Nat. Rev. Neurosci.* **5**, 495–501 (2004).
9. Aks, D. J. and Enns, J. T. *Visual search for size is influenced by a background texture gradient*. *J. Exp. Psychol. Hum. Percept. Perform.* **22**, 1467–1481 (1996).
10. Anderson, B. A., Laurent, P. A. and Yantis, S. *Value-driven attentional capture*. *Proc. Natl. Acad. Sci. USA* **108**, 10367–10371 (2011).
11. Anderson, B. A. and Yantis, S. *Persistence of value-driven attentional capture*. *J. Exp. Psychol. Hum. Percept. Perform.* **39**, 6–9 (2013).
12. Wolfe, J. M., Friedman-Hill, S. R., Stewart, M. I. and O’Connell, K. *The role of categorization in visual search for orientation*. *J. Exp. Psychol. Hum. Percept. Perform.* **18**, 34–49 (1992).
13. Ehinger, K. A., Hidalgo-Sotelo, B., Torralba, A. and Oliva, A. *Modelling search for people in 900 scenes: a combined source model of eye guidance*. *Vis. Cogn.* **17**, 945–978 (2009).
14. Henderson, J. M. and Ferreira, F. *Scene perception for psycholinguists*. In: The Interface of Language, Vision, and Action: Eye Movements and the Visual World. Henderson, J. M. and Ferreira, F., Eds. Psychology Press (2004).
15. Biederman, I., Mezzanotte, R. J. and Rabinowitz, J. C. *Scene perception: detecting and judging objects*

- undergoing relational violations. *Cogn. Psychol.* **14**, 143–177 (1982).
16. Kundel, H. L. and La Follette, P. S. Jr. *Visual search patterns and experience with radiological images.* *Radiology* **103**, 523–528 (1972).
 17. Wooding, D. S., Roberts, G. M. and Phillips-Hughes, J. *The development of the eye-movement response in the trainee radiologist.* *Proc. SPIE* **3663**, 136–145 (1999).
 18. Yantis, S. and Jonides, J. *Attentional capture by abrupt onsets: new perceptual objects or visual masking.* *J. Exp. Psychol. Hum. Percept. Perform.* **22**, 1505–1513 (1996).
 19. Franconeri, S. L., Hollingworth, A. and Simons, D. J. *Do new objects capture attention?* *Psychol. Sci.* **16**, 275–281 (2005).
 20. Sully, J. *The Human Mind: A Text-Book of Psychology.* D. Appleton & Co. (1892).
 21. Wolfe, J. M., Cave, K. R. and Franzel, S. L. *Guided Search: an alternative to the feature integration model for visual search.* *J. Exp. Psychol. Hum. Percept. Perform.* **15**, 419–433 (1989).
 22. Wolfe, J. M. *Guided Search 2.0: a revised model of visual search.* *Psychon. Bull. Rev.* **1**, 202–238 (1994).
 23. Wolfe, J. M. *Guided Search 4.0: current progress with a model of visual search.* In Gray, W. (Ed.) *Integrated Models of Cognitive Systems.* Oxford University Press (2007).
 24. Serences, J. T. and Yantis, S. *Selective visual attention and perceptual coherence.* *Trends Cogn. Sci.* **10**, 38–45 (2006).
 25. Wolfe, J. M. *Moving towards solutions to some enduring controversies in visual search.* *Trends Cogn. Sci.* **7**, 70–76 (2003).
 26. Najemnik, J. and Geisler, W. S. *Optimal eye movement strategies in visual search.* *Nature* **434**, 387–391 (2005).
 27. Renninger, L. W., Verghese, P. and Coughlan, J. *Where to look next? Eye movements reduce local uncertainty.* *J. Vis.* **7**, 6 (2007).
 28. Klein, R. *Inhibitory tagging system facilitates visual search.* *Nature* **334**, 430–431 (1988).
 29. Horowitz, T. S. and Wolfe, J. M. *Visual search has no memory.* *Nature* **394**, 575–577 (1998).
 30. Klein, R. M. and MacInnes, W. J. *Inhibition of return is a foraging facilitator in visual search.* *Psychol. Sci.* **10**, 346–253 (1999).
 31. Wolfe, J. M. *Saved by a log: how do humans perform hybrid visual and memory search?* *Psychol. Sci.* **23**, 698–703 (2012).
 32. Eckstein, M. P. *Visual search: a retrospective.* *J. Vis.* **11**(5), 14 (2011).
 33. Chan, L. K. H. and Hayward, W. G. *Visual search.* Wiley Interdisp. Rev. Cogn. Sci. **4**, 415–429 (2013).
 34. Wolfe, J. M. *Theoretical and behavioral aspects of selective attention.* In: *The Cognitive Neurosciences*, fifth edn. Gazzaniga, M. S. and Mangun, G. R., Eds. MIT Press (2014).
 35. Wolfe, J. M. *Visual search.* In: *The Handbook of Attention.* Fawcett, J., Risko, E. and Kingstone, A., Eds. MIT Press (2014).
 36. Drew, T., Vo, M. L. and Wolfe, J. M. *The invisible gorilla strikes again: sustained inattention blindness in expert observers.* *Psychol. Sci.* **24**, 1848–1853 (2013).
 37. Mack, A. and Rock, I. *Inattention Blindness.* MIT Press (1998).
 38. Simons, D. J. and Chabris, C. F. *Gorillas in our midst: sustained inattention blindness for dynamic events.* *Perception* **28**, 1059–1074 (1999).
 39. Most, S. B., Scholl, B. J., Clifford, E. R. and Simons, D. J. *What you see is what you set: sustained inattention blindness and the capture of awareness.* *Psychol. Rev.* **112**, 217–242 (2005).
 40. Potchen, E. J. *Measuring observer performance in chest radiology: some experiences.* *J. Am. Coll. Radiol.* **3**, 423–432 (2006).
 41. Treisman, A. and Souther, J. *Search asymmetry: a diagnostic for preattentive processing of separable features.* *J. Exp. Psychol. Gen.* **114**, 285–310 (1985).
 42. Lum, T. E., Fairbanks, R. J., Pennington, E. C. and Zwemer, F. L. *Profiles in patient safety: misplaced femoral line guidewire and multiple failures to detect the foreign body on chest radiography.* *Acad. Emerg. Med.* **12**, 658–662 (2005).
 43. Henderson, J. M., Weeks, P. A. and Hollingworth, A. *The effects of semantic consistency on eye movements during complex scene viewing.* *J. Exp. Psychol. Hum. Percept. Perform.* **25**, 210–228 (1999).
 44. Kundel, H. L., Nodine, C. F. and Carmody, D. *Visual scanning, pattern recognition and decision-making in pulmonary nodule detection.* *Invest. Radiol.* **13**, 175–181 (1978).
 45. Nodine, C. F. and Kundel, H. L. *Using eye movements to study visual search and to improve tumor detection.* *Radiographics* **7**, 1241–1250 (1987).
 46. Berbaum, K. S., Brandser, E. A., Franken, E. A., Dorfman, D. D., Caldwell, R. T. and Krupinski, E. A. *Gaze dwell times on acute trauma injuries missed because of satisfaction of search.* *Acad. Radiol.* **8**, 304–314 (2001).
 47. Eckstein, M. P., Beutter, B. R. and Stone, L. S. *Quantifying the performance limits of human saccadic targeting during visual search.* *Perception* **30**, 1389–1401 (2001).
 48. Mello-Thoms, C. *Perception of breast cancer: eye-position analysis of mammogram interpretation.* *Acad. Radiol.* **10**, 4–12 (2003).
 49. Drew, T., Vo, M. L., Olwal, A., Jacobson, F., Seltzer, S. E. and Wolfe, J. M. *Scanners and drillers: characterizing expert visual search through volumetric images.* *J. Vis.* **13**(10), 3 (2013).
 50. Aizenman, A., Thompson, M., Ehinger, K. and Wolfe, J. *Visual search through a 3D volume: studying novices in order to help radiologists.* *J. Vis.* **15**(12), 1107 (2015).
 51. Kundel, H. L. and Nodine, C. F. *Interpreting chest radiographs without visual search.* *Radiology* **116**, 527–532 (1975).
 52. Kundel, H. L., Nodine, C. F., Krupinski, E. A. and Mello-Thoms, C. *Using gaze-tracking data and mixture distribution analysis to support a holistic model for the detection of cancers on mammograms.* *Acad. Radiol.* **15**, 881–886 (2008).
 53. Greene, M. R. and Oliva, A. *The briefest of glances: the time course of natural scene understanding.* *Psychol. Sci.* **20**, 464–472 (2009).
 54. Evans, K. K., Georgian-Smith, D., Tambouret, R., Birdwell, R. L. and Wolfe, J. M. *The gist of the abnormal: above-chance medical decision making in the blink of an eye.* *Psychon. Bull. Rev.* **20**, 1170–1175 (2013).

55. Gur, D., Sumkin, J. H., Rockette, H. E., Ganott, M., Hakim, C., Hardesty, L., Poller, W. R., Shah, R. and Wallace, L. *Changes in breast cancer detection and mammography recall rates after the introduction of a computer-aided detection system.* J. Natl. Cancer Inst. **96**, 185–190 (2004).
56. Mackworth, J. *Vigilance and Attention.* Penguin Books (1970).
57. Wolfe, J. M., Horowitz, T. S. and Kenner, N. M. *Rare targets are often missed in visual search.* Nature **435**, 439–440 (2005).
58. Evans, K. K., Birdwell, R. L. and Wolfe, J. M. *If you don't find it often, you often don't find it: why some cancers are missed in breast cancer screening.* PLoS One. **8**(5), e64366 (2013).
59. Wolfe, J. M. and Van Wert, M. J. *Varying target prevalence reveals two dissociable decision criteria in visual search.* Curr. Biol. **20**, 121–124 (2010).
60. Wolfe, J. M., Horowitz, T. S., Van Mert, M. J., Kenner, N. M., Place, S. S. and Kibbi, N. *Low target prevalence is a stubborn source of errors in visual search tasks.* J. Exp. Psychol. Gen. **136**, 623–638 (2007).
61. Evans, K. K., Tambouret, R. H., Evered, A., Wilbur, D. C. and Wolfe, J. M. *Prevalence of abnormalities influences cytologists' error rates in screening for cervical cancer.* Arch. Pathol. Lab. Med. **135**, 1557–1560 (2011).
62. Wolfe, J. M., Brunelli, D. N., Rubinstein, J. and Horowitz, T. S. *Prevalence effects in newly trained airport checkpoint screeners: trained observers miss rare targets, too.* J. Vis. **13**(3), 1–9 (2013).
63. Berbaum, K. S. *et al.* *Satisfaction of search in diagnostic radiology.* Invest. Radiol. **25**, 133–140 (1990).
64. Cain, M. S., Biggs, A. T., Darling, E. F. and Mitroff, S. R. *A little bit of history repeating: splitting up multiple-target visual searches decreases second-target miss errors.* J. Exp. Psychol. Appl. **20**, 112–125 (2014).