Published in final edited form as:

J Am Coll Radiol. 2016 May ; 13(5): 571–575. doi:10.1016/j.jacr.2015.11.023.

Let's Use Cognitive Science to Create Collaborative Workstations

Murray Reicher, MD, FACR and

Chief Medical Officer, Merge Healthcare, Inc

Jeremy M Wolfe

Brigham and Women's Hospital/Harvard Medical School

Radiologists are drowning in images and health information. Suppose a radiologist views 50 CT, MRI, and/or PET exams in a day. Each of those exams might generate 1000 images (2000 with one prior exam for comparison). That adds up to roughly 100,000 images per day. In addition, while reading, a radiologist must take into account the patient's electronic medical record and exploding scientific literature. Our human limitations are such that only a subset of those images can get real scrutiny and only a fraction of the medical record and literature can be weighed. The stress of an insurmountable job is contributing to radiology burn-out.¹ It's time for technology that can help radiologists thrive in an era of information overload.

Computer-aided detection and diagnosis could certainly help. Today, computers in radiology can help detect breast lesions and lung nodules, analyze breast, liver, and brain contrast enhancement patterns, assess vascular stenoses, detect cardiac wall motion abnormalities, and more. Furthermore, research suggests that computers may accurately differentiate normal vs. abnormal chest radiographs.² Perhaps soon, computers will triage images within an exam or entire exams to determine which require physician evaluation.

While we hope for further advances in computer-aided detection, our purpose here is to discuss a complementary route to enhancing the life of the radiologists. We have labeled this approach "Perceptual Design". The goal of perceptual design is to use a scientific understanding of human perception and cognition to guide how images and information are presented to and analyzed by radiologists. This approach can provide low-cost innovations that take better advantage of our human strengths and compensate for our weaknesses.

In what follows, we will sketch several examples of perceptual design, starting with ideas that have some research to back them up and ending with more speculative possibilities. In each case, we discuss how design informed by our understanding of human perception, cognition, bias, focus, and emotion can result in better job performance and satisfaction.

Toggling between images

Until the last two decades, doctors compared medical images by positioning them side-by-side because those images were physical pieces of film. The advent of PACS and computerized workstations made alternatives possible. For instance, using computerized image display, it became possible to present a sequence of still images (e.g. slices in a chest

CT) as a movie.³ Paging through anatomically ordered cross-sectional medical images (“stack mode”) has now been commonly adopted, largely replacing the side-by-side comparison of cross-sectional images as was required by film.

While stack mode reading has become common for cross-sectional imaging, when it comes to comparing images from two different exams or two series in the same exam, the most common clinical practice is still side-by-side comparison as required in the old film-based paradigm. Perceptual science suggests that comparison of medical images by toggling or flickering between comparable images in the same screen location may result in better human performance. Instead of viewing images side by side, the comparable images from new and prior exams are arranged in an interleaved stack and the reader moves between the current and older exams in time rather than in space. Such a technique was used to find Pluto⁴, and is used today for a wide range of imaging applications including comparing retinal photographs and gel electrophoreses.^{5–8}

Unlike astronomy, in mammography, last year’s and the year’s images cannot be perfectly aligned. Nevertheless, when non-experts compare photographs⁹ or radiologists compare mammograms¹⁰, observers are significantly faster to find a change when images are toggled than when they are viewed side-by-side. This can probably be attributed to a reduction in the number of large saccadic eye movements that are needed in the toggle condition. In side-by-side viewing, any interesting feature in one image must be checked in the other image in a physically different location. In the toggle condition, that interesting feature can be examined without moving the eyes and without the possibility that the large eye movement might not land in exactly the right spot. The data on toggling while reading mammograms also suggest an improvement in radiologist accuracy that needs further evaluation.¹⁰ Even if it only produced a time savings, toggling images for purposes of comparison could be an example of a simple, low-cost perceptual design-based intervention that produces real benefit. As automated registration of image position and magnification improve, the benefits of toggling could be further enhanced. For example, with image volumes such as produced by CT, MRI, and PET, three dimensional registration can play a role in creating more perfectly matching comparison images.

Perceptual design to address the prevalence problem in screening

In situations like breast cancer screening, radiologists are looking for rare targets (a fraction of 1% of a screening population).¹¹ Unfortunately, humans are not built to search successfully for low probability items. Since World War II, research has shown that observers suffer a “vigilance decrement”^{12–14} when monitoring a display for a signal like a ‘blip’ on a radar screen. When viewing a stack of image, observers are also more likely to miss a target item if it is rare than if it is more common.¹⁵ The same cancers can be missed when encountered in the course of regular, low prevalence screening but found when encountered in an enriched, high prevalence stack of cases.¹⁶ This is not a simple decline in accuracy.¹⁷ It is not that low prevalence makes radiologists sloppy. Rather, knowledge of the likelihood of finding a lesion affects the radiologist’s performance. At low prevalence, false negative errors go up and false positives go down. At high prevalence, false positives go up and false negatives go down.¹⁸ This is the mark of a criterion shift and is a general problem,

showing up, for example, when cytologists screen for cervical cancer in Pap tests slides¹⁹ or when transportation security officers look for threats in carry-on luggage.²⁰

Low prevalence presents a special challenge for diagnostic tasks like breast cancer screening. A clinician, who reads only a modest number of mammography cases (e.g. in a general radiology practice) is not going to see many positive cases in the course of a year. This probably reduces both sensitivity and specificity.²¹ Various solutions have been suggested; for example, adding cases with known pathology into the workflow.¹⁶ It is probably not practical to add enough cases to shift criterion as much as might be desired because of the increase in caseload that would be required.²² However, the added cases would, at least, expose readers to more positive examples. The use of “test sets”, where radiologists read dozens of cases as a form of CME, can serve a similar purpose of “tuning up” the reader’s internal template of what cancer looks like.²¹ Test sets, where implemented, tend to be read once a year. In the realm of perceptual design, an even better approach might be to have a much briefer task, performed prior to a screening session, in which the radiologist quickly evaluates a relatively small number of images, receiving instant feedback. With a bit of effort, this might be turned into a game wherein individuals would compete with themselves and each other to help turn this perceptual tune-up into an appealing part of the daily or weekly routine.

What can be learned from eye-tracking?

Researchers in medical image perception have been looking at the eye movements of radiologists for decades.²³ Radiologists make fewer eye movements over an image as they become more expert,²⁴ having learned where *not* to look. Eye movement research enables classification of mistakes into “search errors” - where the target was never fixated (eyes directed at the target), “recognition errors” – where the target was fixated but only briefly, and “decision errors” – where the target is scrutinized but not correctly reported.^{25,27} More recent studies have assessed eye movements as radiologists scroll through 3D volumes of images as in lung CT²⁶ or breast tomosynthesis.²⁷ When reading lung CT, readers fell into two types. There were “drillers”, who kept their eyes in a relatively limited area of the 2D image while drilling through the depth of the stack of images. They would then move their eyes to another location in the 2D plane and drill again. “Scanners”, in contrast, would move only slowly in depth while scanning all over the 2D image. While there is as yet insufficient data to conclude which method is better, the answer would have implications for how radiologists are trained to search stacks of images. For example, a workstation cognizant of eye movements could automatically adjust to each reader’s ideal paging speed when riffling through a stack of images.

Moreover, as eye tracking becomes less intrusive, it could become an integrated part of future workstations. For example, when the radiologist indicates that it is time to move to the next case, an eye-tracking workstation could let the radiologist see what portions of an image or stack of image were never examined. Perhaps a workstation cognizant of eye movements could be programmed to automatically magnify or otherwise transform a region, if that region were persistently fixated (akin to tying computer actions to the position of the mouse, hovering over an advertisement, for example). A workstation that monitored the

oculomotor behavior of a radiologist, might be able to learn hanging protocols specifically suited to that radiologist.

Perceptual design, job satisfaction, and efficiency

In his classic description of the mental state of flow, psychologist Mihály Csíkszentmihályi describes what he calls “flow,” an enjoyable state of concentration that leads to increased performance, further learning, and skill development.”²⁸ Flow comes from confronting challenging, but achievable tasks with clear goals. To achieve flow, the user must be able to focus; to feel a sense of control and mastery. Is that still achievable given the radiologist’s progressively pressured information processing burden? Perhaps workstations designed based on an understanding of human perception and cognition can help by becoming the radiologist’s collaborator or assistant.

For example, software might display clear goals and provide feedback as the user makes progress. Automated workflow management software might automatically assign exams based on preferences or skill attributes. Integrated hardware/software design can also play a role: a workstation might learn to bring the most relevant content to the radiologist’s central visual field, reducing the need to scan multiple displays for information.

Interruptions will happen, but the workstation could serve as the personal assistant who protects the radiologist’s focus. For example, a messaging system built into PACS might only interrupt a radiologist during the transition between exams. Even more cleverly, a system might selectively route a technologist’s call to the radiologist who has been reading non-stop for over an hour, thus intelligently meting out a break to the physician most in need of one. The system might monitor the radiologist’s computer activities on order to detect and warn about signs of fatigue.

The collaborative workstation could learn to provide a customized set of tools based on reader preferences, the displayed exam type, reader specialty, and prior habits. Instead of a traditional mouse and keyboard as input devices, a tablet could double as an input device and communication assistant, not only presenting personalized tools, but also providing the reader with context sensitive training tutorials, service chats, reference materials, protocols, and phone/text contacts. The result could be a greater sense of mastery and control.

A collaborative workstation can also promote flow by progressively adapting as the user’s skill increases. A video game quickly becomes boring and unrewarding if it fails to change as user skill advances, yet with the typical clinical reporting/PACS workstation, the user interface remains the same for a beginner as for an advanced user. For example, a user who has learned an advanced skill, such as how to access common functions via a keyboard shortcut, audio command, or other tool access shortcut is today often left with a screen littered with no-longer-employed toolbars and buttons. By adopting designs that evolve to match the user’s mastery, work could be made more enjoyable while performance is improved.

The collaborative workstation can compile big data from electronic medical records and the medical literature to assist in determining probabilities of various possible diagnoses. Rather

than presenting the radiologist with an overwhelming amount of clinical information from disparate sources, imagine instead that the radiologist is shown a compiled list of relevant clinical information and statistically likely new diagnoses. An analogy is the back of a baseball card. At a glance, an individual with domain expertise (i.e., a familiarity with the statistics of the game) can glance at the back of a baseball card and nearly instantly infer many narratives about a player—age, position, experience, success, playing attributes, and more. Whereas a more detailed history of every event in a player's career may be counterproductive, the back of the card tells the reader just enough for him to decide whether or not he wishes to research more deeply and learn more.

Perceptual design can also help build better reading habits. For example, some radiologists customarily page through images while dictating. Others assemble a montage of their key selected images and only then report the results. Some view the report constantly while using speech recognition. Others have the computer provide text-to-speech prompts to specify the report location via audio, then only proof the report at the end of the reading process. If this information was collected and compiled (with appropriate protection for confidentiality), some techniques may be found to produce more efficient or accurate reading. Then, the collaborative workstation could help inform readers of their habits in order to help promote voluntary change. Benchmarking tools might also inform the user of their number of clicks per read, the mouse drag distance per read, the number of key images selected per read, the frequency at which a critical result is flagged, the number or total time of calls or other interactions with referring physicians, the number of exams marked for teaching purposes, of reports generated with evidence-based recommendations, the number of practice-improvement suggestions submitted, or other metrics that provide a user with personal measures of self-improvement.

If the intuition and expertise of radiologists is partnered with the tools of perceptual and cognitive science, significant advances can be made. These advances will be different from the development of a new image reconstruction algorithm or a new imaging modality. However, just as we should not stop seeking new technological breakthroughs, we should not stop improving the ways in which humans interact with that technology.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Moskowitz, Peter. Gathering storm clouds suggest the need for a change in radiology. Radiologist-centered Imaging. Radiology. 2015; 276:3–7. [PubMed: 25989293]
2. Jaeger S, Karargyris A, Candemir S, Folio L, Siegelman J, Callaghan F, Zhiyun Xue, Palaniappan K, Singh RK, Antani S, Thoma G, Yi-Xiang Wang, Pu-Xuan Lu, McDonald CJ. Automatic tuberculosis screening using chest radiographs. IEEE Trans Med Imaging. 2014 Feb; 33(2):233–45. [PubMed: 24108713]
3. Seltzer SE, Judy PF, Adams DF, Jacobson FL, Stark P, Kikinis R, et al. Spiral CT of the chest: comparison of cine and film-based viewing. Radiology. 1995; 197(1):73–78. [PubMed: 7568857]
4. Tombaugh CW. The Search for the Ninth Planet, Pluto. Astronomical Society of the Pacific Leaflets. 1946; 5(209):73.

5. Lemkin, PF., Thornwall, G., Evans, J. Comparing 2-D Electrophoretic Gels Across Internet Databases. In: Walker, John, editor. The Protein Protocols Handbook. Humana Press Inc; Totowa, NJ: 2005. p. 279-305.
6. VanderBeek BL, Smith SD, Radcliffe NM. Comparing the detection and agreement of parapapillary atrophy progression using digital optic disk photographs and alternation flicker. Graefes Arch Clin Exp Ophthalmol. 2010 Sep; 248(9):1313–7. [PubMed: 20393740]
7. Funk J, Lagrèze W, Zeyen T. European Glaucoma Prevention Study Group. Flicker comparison of optic disc photographs: sensitivity and specificity. Klin Monbl Augenheilkd. 2002 Dec; 219(12): 862–5. [PubMed: 12548470]
8. Berger JW, Patel TR, Shin DS, Piltz JR, Stone RA. Computerized stereochronoscopy and alternation flicker to detect optic nerve head contour change. Ophthalmology. 2000 Jul; 107(7):1316–20. [PubMed: 10889105]
9. Josephs EL, Drew T, Wolfe JM. Shuffling your way out of change blindness. Psych Bulletin & Review. 2015 in press.
10. Drew T, Aizenman AM, Thompson MB, Kovacs MD, Trambert M, Reicher MA, et al. Image toggling saves time in mammography. J Medical Imaging. 2015 submitted.
11. Gur D, Sumkin JH, Rockette HE, Ganott M, Hakim C, Hardesty L, et al. Changes in breast cancer detection and mammography recall rates after the introduction of a computer-aided detection system. J Natl Cancer Inst. 2004; 96(3):185–190. [PubMed: 14759985]
12. Colquhoun WP, Baddeley AD. Influence of signal probability during pretraining on vigilance decrement. J Exp Psychol. 1967; 73(1):153–155. [PubMed: 6031659]
13. See JE, Howe SR, Warm JS, Dember WN. Meta-Analysis of the Sensitivity Decrement in Vigilance. Psychological Bulletin. 1995; 117(2):230–249.
14. Swets JA, Kristofferson AB. Attention. Annu Rev Psychol. 1970; 21:339–366. [PubMed: 4904374]
15. Wolfe JM, Horowitz TS, Kenner NM. Rare targets are often missed in visual search. Nature. 2005; 435(7041):439–440. DOI: 10.1038/435439a [PubMed: 15917795]
16. Evans KK, Birdwell RL, Wolfe JM. If You Don't Find It Often, You Often Don't Find It: Why Some Cancers Are Missed in Breast Cancer Screening. PLoS ONE. 2013; 8(5):e64366.doi: 10.1371/journal.pone.0064366 [PubMed: 23737980]
17. Gur D, Bandos AI, Cohen CS, Hakim CM, Hardesty LA, Ganott MA, et al. The “Laboratory” Effect: Comparing Radiologists’ Performance and Variability during Prospective Clinical and Laboratory Mammography Interpretations. Radiology. 2008; 249(1):47–53. [PubMed: 18682584]
18. Wolfe JM, Van Wert MJ. Varying Target Prevalence Reveals Two Dissociable Decision Criteria in Visual Search. Curr Biol. 2010; 20(2):121–124. S0960-9822(09)02122-8 [pii]. [PubMed: 20079642]
19. Evans KK, Tambouret R, Wilbur DC, Evered A, Wolfe JM. Prevalence of Abnormalities Influences Cytologists’ Error Rates in Screening for Cervical Cancer. Archives of Pathology & Laboratory Medicine. 2011; 135(12):1557–1560. DOI: 10.5858/arpa.2010-0739-OA [PubMed: 22129183]
20. Wolfe JM, Brunelli DN, Rubinstein J, Horowitz TS. Prevalence effects in newly trained airport checkpoint screeners: trained observers miss rare targets, too. J of Vision. 2013; 13(3):33.doi: 10.1167/13.3.33
21. Nass, SJ., Patlak, Margie. National Academies of Sciences, E., and Medicine. Assessing and Improving the Interpretation of Breast Images: Workshop Summary. 2015.
22. Wolfe JM, VanWert MJ. Varying target prevalence reveals two, dissociable decision criteria in visual search. Curr Biol. 2010; 20(2):121–124. DOI: 10.1016/j.cub.2009.11.066 [PubMed: 20079642]
23. Kundel HL, Nodine CF, Carmody D. Visual scanning, pattern recognition and decision-making in pulmonary nodule detection. Invest Radiol. 1978; 13(3):175–181. [PubMed: 711391]
24. Kundel HL, La Follette PS Jr. Visual search patterns and experience with radiological images. Radiology. 1972; 103(3):523–528. [PubMed: 5022947]
25. Mello-Thoms C. Perception of breast cancer: eye-position analysis of mammogram interpretation. Acad Radiol. 2003; 10(1):4–12. [PubMed: 12529023]

26. Drew T, Vo ML-H, Olwal A, Jacobson F, Seltzer SE, Wolfe JM. Scanners and drillers: Characterizing expert visual search through volumetric images. *Journal of Vision*. 2013; 13(10)doi: 10.1167/13.10.3
27. Aizenman, A., Drew, T., Georgian-Smith, D., Wolfe, JM. Comparing search strategy in breast tomosynthesis and 2D mammogram: an eye tracking study; paper presented at the Annual Meeting of the Vision Science Society; May 2014; 2014. p. 418
28. Csíkszentmihályi, M. *Flow: The Psychology of Optimal Experience* (Harper Perennial Modern Classics) Paperback. 2008.