

The role and benefits of cognition and perception in medical imaging

This manuscript ([permalink](#)) was automatically generated from [thorowitz/ThinkTankClinicians@5fd9772](#) on April 22, 2020.

Authors

- **Melissa Treviño**

National Cancer Institute

- **Todd S. Horowitz**

 [0000-0001-5571-6266](#) ·  [thorowitz](#)

National Cancer Institute

Introduction

In an era of unprecedented technological advances across the field of medical imaging, it is critically important to understand the cognitive and perceptual functioning of the “human in the loop,” the clinicians who make the final decisions on detection, diagnosis, and treatment. However, research on the cognitive and perceptual processes underlying clinician performance is chronically understudied. In September 2019, the National Cancer Institute (NCI) convened the “Cognition and Medical Image Perception Think Tank” in order to advance research on medical image perception¹ and cognition. The Think Tank brought radiologists and pathologists together with researchers working in medical image perception and adjacent fields of cognition and perception, along with representatives from interested federal government agencies (including NCI, National Institute of Biomedical Imaging and Bioengineering, U.S. Food and Drug Administration, U.S. Department of Homeland Security) to address the Think Tank’s key objectives (see Table ??) through panel discussions, constructive discourse, idea generation, and problem-solving. Panel discussions were guided by critical questions generated by the participants themselves in advance of the meeting. Here, we report on the Think Tank’s deliberations and highlight opportunities to expand the research in medical image perception and cognition.

Objectives
1. Identify research gaps and critical unsolved problems in pathology and radiology, from the perspective of clinicians.
2. Discuss ways to address these clinically-informed questions through cognitive and perceptual research.
3. Identify barriers that hinder collaborations between researchers and clinicians and identify potential solutions.
4. Identify and discuss ways to elevate psychology’s profile within the medical image community and broader communities.
5. Identify future goals and identify strategies to evaluate progress.

Table 1. Think Tank Objectives {#tbl:Objectives}

Identifying research gaps

Detection and diagnosis rely on clinicians² searching for abnormalities in medical images. Across the many ways to acquire images (from light microscopy to magnetic resonance) and present them (e.g., via glass slides or computer workstations), clinicians are confronted with similar perceptual and cognitive challenges. Clinicians are tasked with bringing to bear a vast amount of medical knowledge, together with contextual information about the patient (e.g., referring physician’s information, prior report, prior medical history, genomics information) in order to interpret highly complex and variable images quickly and accurately. This is a highly challenging endeavor. Clinicians face many difficulties, from information overload to repeated interruptions to clinician burnout, that in turn are closely associated with health care quality, patient safety, and financial burdens on practices [???,^{1,2}]. Even under ideal circumstances, human observers are bound to make mistakes. Errors, both false positives and false negatives, are a reality in both radiology [^{3,4,5}] and pathology practice [⁶]. The majority of diagnostic errors in radiology can be traced to perceptual and cognitive factors (Berlin & Berlin, 1995; Smith, 1967). The same is probably true in pathology, though there is much less research in this area (Raab et al., 2005). Improving the accuracy of diagnosis requires an understanding of the perceptual and cognitive mechanisms underlying the performance of clinicians. Additionally, cognitive and perceptual principles and approaches should be applied to address issues concerning clinicians’ work environment and well-being.

What research should we encourage in order to improve clinicians' diagnostic accuracy, work environment, and well-being? The first challenge for the Think Tank participants was to identify the research gaps and critical unsolved problems in pathology and radiology, from the perspective of clinicians. The primary challenges identified by the clinicians were information overload and integrating large amounts of information from various sources.

For example, the transition from the analog to digital eras has meant a qualitative change in the nature of the radiological problem. A radiologist scrutinizing a single chest X-ray for a lung nodule is trying to extract a target with a very low signal:noise ratio (SNR) from a single image. The same radiologist reading a chest CT is searching for a target with a fairly high SNR in a scan comprising thousands of images (slices). In addition, the radiologist must now incorporate into her decision information from outside sources, such as electronic health records and genetic panels, that may not have been available in the past. Furthermore, the decisions that clinicians have to make have gotten more complex as precision medicine has advanced. A pathologist now might have to indicate not just whether something is a tumor or not, but if the tumor is hot or cold, likely to be responsive or non-responsive. The cognitive capacity of the clinician, in contrast, has not changed, and the time pressure has only gotten worse. What are the effects of information overload on medical decision making and decision quality? Are there technical solutions or decision-making aids or strategies that could be used to efficiently integrate complex information? For example, psychologically distancing oneself from information has been shown to improve non-medical decision-making (Fukukura et al., 2013). What is the applicability of decision strategies to real world practice? The cognitive and perceptual challenges of the clinicians' role, and the implications for practice, are poorly understood at present.

Task shifting was another topic identified by the Think Tank participants. Interruptions and distractions from a variety of sources (e.g., texts, email, pager, colleagues, telephone) may lead clinicians to miss or fail to report important findings. Another form of task shifting is when clinicians move from case to case, different organs, while using various image modalities. There is no research on whether task shifting in medical imaging creates adverse sequential consequences and/or how to mitigate these effects.

As one might expect from the above descriptions, the workload for clinicians has increased, and physical and mental fatigue of clinicians are at unprecedented levels, contributing to both perceptual and interpretive errors. Fatigue can manifest as attention errors, such as when a clinician scrolls too fast through multi-slice or volumetric images, which may give the false impression that the clinician has sufficiently viewed the case. Additional errors can result from lack of focus, getting easily distracted, and/or failing to record what is perceived. These errors caused by physical and mental fatigue become more apparent at the end of the day. Multiple studies have reported a decrease in detection accuracy due to fatigue (for a review, see Waite et al., 2017) but translational research needs to be conducted to mitigate fatigue effects in real world practice.

Interpretative errors can occur from a variety of sources. For radiologists, several categories of tumors can be difficult to interpret, such as when there is a subtle gradient (i.e., when it is difficult to notice subtle changes or new lesions) from benign to malignant tumors, malignant to benign tumors, and when benign tumors mimic malignant tumors. Additionally, it is also often difficult to interpret or connect two dissonant findings (e.g., adrenal nodule with a lung mass) and the clinician ends up failing to link findings and report a complete picture. Research should focus on understanding the perceptual and cognitive underpinnings of these type of cases to help mitigate interpretation difficulty.

Artificial intelligence (AI) tools are often proposed as a solution to the challenges facing clinicians. However, there is a mismatch between what clinicians need from AI and how AI is typically developed. AI solutions are often pitched as "outperforming" the clinician, implying that AI systems are on the cusp of replacing the clinician (Mukherjee, 2017), or at least serving as an equivalent reader (McKinney

et al., 2020). This stems from an incomplete picture of the clinician's job. A better approach would be to develop AI as a tool to automate the data reduction process, extracting the information that will aid the clinician's decision process. For example, when reading pap smears, cytotechnologists can sign off on cases that are unequivocally negative, leaving the cytopathologists to focus only on cases with potential abnormalities. AI could usefully fulfill this sort of role, and potentially automate quality assurance (QA) procedures. A useful framework is the "diagnostic cockpit of the future", described by Krupinski et al. (2019) as a set of tools to "aggregate, organize, and simplify medical imaging results and patient-centric clinical data." This metaphor underscores the role of AI in helping to winnow the information available to the clinician and present only the most clinically relevant information in the clearest way possible. In order to accomplish these goals, AI should be developed within a broad interdisciplinary context. Clinical, perceptual, and cognitive perspectives are all needed in the development process, rather than treated as an afterthought.

Addressing clinical questions through perceptual research

There's an old science joke that goes something like this: A dairy farmer is looking for ways to increase the output of his cows, as cheaply as possible. He consults an agronomist, who recommends changes to the feedstock and new equipment which will require major investments. He turns to a biologist, who suggests hormone treatments that will be cheaper, but still involve an outlay of more money than he wants to spend. Finally, he turns to a physicist, who promises a free solution. He eagerly asks her to explain her idea. She begins, "First, assume a spherical cow with radius r and mass m , uniformly emitting milk in all directions..."

To many clinicians, perceptual research on medical image interpretation can sometimes seem like the equivalent of a spherical cow emitting milk in all directions. Perceptual studies typically abstract away from the specificity and messiness of the clinical reading context in order to gain experimental control. Complex stimuli are simplified, narrative reports are replaced with binary responses, clinical history is absent, and the life-and-death stakes of the clinic are removed. Consider Wolfe and Horowitz's work on the effects of prevalence (Wolfe et al., 2007), in which amateur observers drawn from the Cambridge, MA community searching for the letter "T" concealed among arrays of "L"s are meant to have some relevance to radiologists searching for cancerous lesions in mammograms or lung CTs. Traditional medical image perception research may skew closer to the clinical experience, using experienced professional readers and real cases. Even so, the consequences for a missed target or unnecessary callback are not the same. Furthermore, in perceptual research each trial in a given experimental condition is assumed to be equivalent to other trials (because stimuli are typically randomly and algorithmically generated), to the clinician, each case is unique, leading to a reluctance to generalize. All of these factors may lead to a certain skepticism in the clinician community about the value of perception research, especially in the face of the difficulties raised in the *Identifying Research Gaps* section.

There is a trend in medical image perception studies where the researcher follows a "Reverse Translation" model to assess clinical problems. This is where a researcher observes a problem in the clinic that resembles an area of perception research they are doing in the lab or has familiarity with. The researcher will then develop lab versions of the problem to study the basic perceptual and cognitive science principles of the problem. These studies will sometimes use observer models or undergraduate students as they are more efficient to use lab studies as clinicians time is limited. After many iterations of working out the basic science and once the researcher has a good working hypothesis, they will then go back in the clinical setting to test their hypothesis. A major issue with this model is that the problems that are being assessed are not the critical issues that clinicians actually want answers to. We propose researchers adopt a "Reverse Translation Phase 2" model, where researchers investigate problems that are important to the clinician and then follow the same "Reverse Translation" procedure.

Within the “Reverse Translation Phase 2” model, there are key assumptions that allow for the translation of principles studied and derived from lab studies to the clinical setting even when clinicians or medical images are not being utilized. First, clinicians and non-clinicians have the same visual systems and perceptual and cognitive faculties that allow us to process and interpret stimuli. Second, artificial tasks and stimuli capture the key elements of the clinical tasks and stimuli. Going back to the work on prevalence effects, studying the effect of low prevalence on radiologists’ performance with a reader study would take years to complete and would put a substantial burden on clinicians. Medical image perception studies have replicated the prevalence effect in the lab setting using non-clinician readers and non-medical image stimuli (Wolfe et al., 2005), demonstrated low prevalence influences non-clinician and clinician readers’ response criteria (Evans et al., 2011, 2013; Wolfe & Van Wert, 2010), and have proposed potential solutions to mitigate prevalence effects (Lau & Huang, 2010; Navalpakkam et al., 2009; Schwark et al., 2012; Wolfe et al., 2013).

Addressing clinical questions through perceptual research

There’s an old science joke that goes something like this: A dairy farmer is looking for ways to increase the output of his cows, as cheaply as possible. He consults an agronomist, who recommends changes to the feedstock and new equipment which will require major investments. He turns to a biologist, who suggests hormone treatments that will be cheaper, but still involve an outlay of more money than he wants to spend. Finally, he turns to a physicist, who promises a free solution. He eagerly asks her to explain her idea. She begins, “First, assume a spherical cow with radius r and mass m , uniformly emitting milk in all directions...”

To many clinicians, perceptual research on medical image interpretation can sometimes seem like the equivalent of a spherical cow emitting milk in all directions. Perceptual studies typically abstract away from the specificity and messiness of the clinical reading context in order to gain experimental control. Complex stimuli are simplified, narrative reports are replaced with binary responses, clinical history is absent, and the life-and-death stakes of the clinic are removed. Consider Wolfe and Horowitz’s work on the effects of prevalence [7], in which amateur observers drawn from the Cambridge, MA community searching for the letter “T” concealed among arrays of “L”s are meant to have some relevance to radiologists searching for cancerous lesions in mammograms or lung CTs. Traditional medical image perception research may skew closer to the clinical experience, using experienced professional readers and real cases. Even so, the consequences for a missed target or unnecessary callback are not the same. Furthermore, in perceptual research each trial in a given experimental condition is assumed to be equivalent to other trials (because stimuli are typically randomly and algorithmically generated), to the clinician, each case is unique, leading to a reluctance to generalize. All of these factors may lead to a certain skepticism in the clinician community about the value of perception research, especially in the face of the difficulties raised in the *Identifying Research Gaps* section.

There is a trend in medical image perception studies where the researcher follows a “Reverse Translation” model to assess clinical problems. This is where a researcher observes a problem in the clinic that resembles an area of perception research they are doing in the lab or has familiarity with. The researcher will then develop lab versions of the problem to study the basic perceptual and cognitive science principles of the problem. These studies will sometimes use observer models or undergraduate students as they are more efficient to use lab studies as clinicians time is limited. After many iterations of working out the basic science and once the researcher has a good working hypothesis, they will then go back in the clinical setting to test their hypothesis. A major issue with this model is that the problems that are being assessed are not the critical issues that clinicians actually want answers to. We propose researchers adopt a “Reverse Translation Phase 2” model, where researchers investigate problems and that are important to the clinician and then follow the same “Reverse Translation” procedure.

Within the “Reverse Translation Phase 2” model, there are key assumptions that allow for the translation of principles studied and derived from lab studies to the clinical setting even when clinicians or medical images are not being utilized. First, clinicians and non-clinicians have the same visual systems and perceptual and cognitive faculties that allow us to process and interpret stimuli. Second, artificial tasks and stimuli capture the key elements of the clinical tasks and stimuli. Going back to the work on prevalence effects, studying the effect of low prevalence on radiologists’ performance with a reader study would take years to complete and would put a substantial burden on clinicians. Medical image perception studies have replicated the prevalence effect in the lab setting using non-clinician readers and non-medical image stimuli [\[8\]](#), demonstrated low prevalence influences non-clinician and clinician readers’ response criteria [\[9,10,11\]](#), and have proposed potential solutions to mitigate prevalence effects [\[12,13,14,15\]](#).

References

- 1. Association Between Physician Burnout and Patient Safety, Professionalism, and Patient Satisfaction**
Maria Panagioti, Keith Geraghty, Judith Johnson, Anli Zhou, Efharis Panagopoulou, Carolyn Chew-Graham, David Peters, Alexander Hodgkinson, Ruth Riley, Aneez Esmail
JAMA Internal Medicine (2018-10-01) <https://doi.org/dc6c>
DOI: [10.1001/jamainternmed.2018.3713](https://doi.org/10.1001/jamainternmed.2018.3713) · PMID: [30193239](https://pubmed.ncbi.nlm.nih.gov/30193239/) · PMCID: [PMC6233757](https://pubmed.ncbi.nlm.nih.gov/PMC6233757/)
- 2. The Business Case for Investing in Physician Well-being**
Tait Shanafelt, Joel Goh, Christine Sinsky
JAMA Internal Medicine (2017-12-01) <https://doi.org/gc7rs9>
DOI: [10.1001/jamainternmed.2017.4340](https://doi.org/10.1001/jamainternmed.2017.4340) · PMID: [28973070](https://pubmed.ncbi.nlm.nih.gov/28973070/)
- 3. Accuracy of Diagnostic Procedures: Has It Improved Over the Past Five Decades?**
Leonard Berlin
American Journal of Roentgenology (2007-05) <https://doi.org/brzf2j>
DOI: [10.2214/ajr.06.1270](https://doi.org/10.2214/ajr.06.1270) · PMID: [17449754](https://pubmed.ncbi.nlm.nih.gov/17449754/)
- 4. Multiparametric Magnetic Resonance Imaging in the Diagnosis of Prostate Cancer: A Systematic Review**
M. A. Haider, X. Yao, A. Loblaw, A. Finelli
Clinical Oncology (2016-09) <https://doi.org/f8zbrf>
DOI: [10.1016/j.clon.2016.05.003](https://doi.org/10.1016/j.clon.2016.05.003) · PMID: [27256655](https://pubmed.ncbi.nlm.nih.gov/27256655/)
- 5. Clinical utility of transperineal template-guided mapping biopsy of the prostate after negative magnetic resonance imaging-guided transrectal biopsy**
Arjun Sivaraman, Rafael Sanchez-Salas, Hashim U. Ahmed, Eric Barret, Nathalie Cathala, Annick Mombet, Facundo Uriburu Pizarro, Arie Carneiro, Steeve Doizi, Marc Galiano, ... Xavier Cathelineau
Urologic Oncology: Seminars and Original Investigations (2015-07) <https://doi.org/f3gwjn>
DOI: [10.1016/j.urolonc.2015.04.005](https://doi.org/10.1016/j.urolonc.2015.04.005) · PMID: [25957713](https://pubmed.ncbi.nlm.nih.gov/25957713/)
- 6. Clinical impact and frequency of anatomic pathology errors in cancer diagnoses**
Stephen S. Raab, Dana Marie Grzybicki, Janine E. Janosky, Richard J. Zarbo, Frederick A. Meier, Chris Jensen, Stanley J. Geyer
Cancer (2005-11-15) <https://doi.org/bs52m7>
DOI: [10.1002/cncr.21431](https://doi.org/10.1002/cncr.21431) · PMID: [16216029](https://pubmed.ncbi.nlm.nih.gov/16216029/)
- 7. Low target prevalence is a stubborn source of errors in visual search tasks.**
Jeremy M. Wolfe, Todd S. Horowitz, Michael J. Van Wert, Naomi M. Kenner, Skyler S. Place, Nour Kibbi
Journal of Experimental Psychology: General (2007) <https://doi.org/fv56d2>
DOI: [10.1037/0096-3445.136.4.623](https://doi.org/10.1037/0096-3445.136.4.623) · PMID: [17999575](https://pubmed.ncbi.nlm.nih.gov/17999575/) · PMCID: [PMC2662480](https://pubmed.ncbi.nlm.nih.gov/PMC2662480/)
- 8. Rare items often missed in visual searches**
Jeremy M. Wolfe, Todd S. Horowitz, Naomi M. Kenner
Nature (2005-05-25) <https://doi.org/bzvfb4>
DOI: [10.1038/435439a](https://doi.org/10.1038/435439a) · PMID: [15917795](https://pubmed.ncbi.nlm.nih.gov/15917795/) · PMCID: [PMC4224304](https://pubmed.ncbi.nlm.nih.gov/PMC4224304/)
- 9. If You Don't Find It Often, You Often Don't Find It: Why Some Cancers Are Missed in Breast Cancer Screening**
Karla K. Evans, Robyn L. Birdwell, Jeremy M. Wolfe

10. Prevalence of Abnormalities Influences Cytologists' Error Rates in Screening for Cervical Cancer

Karla K Evans, Rosemary H Tambouret, Andrew Evered, David C Wilbur, Jeremy M Wolfe

Archives of Pathology & Laboratory Medicine (2011-12) <https://doi.org/dsnc9h>

DOI: [10.5858/arpa.2010-0739-oa](https://doi.org/10.5858/arpa.2010-0739-oa) · PMID: [22129183](https://pubmed.ncbi.nlm.nih.gov/22129183/) · PMCID: [PMC3966132](https://pubmed.ncbi.nlm.nih.gov/PMC3966132/)

11. Varying Target Prevalence Reveals Two Dissociable Decision Criteria in Visual Search

Jeremy M. Wolfe, Michael J. Van Wert

Current Biology (2010-01) <https://doi.org/fcf5fw>

DOI: [10.1016/j.cub.2009.11.066](https://doi.org/10.1016/j.cub.2009.11.066) · PMID: [20079642](https://pubmed.ncbi.nlm.nih.gov/20079642/) · PMCID: [PMC2818748](https://pubmed.ncbi.nlm.nih.gov/PMC2818748/)

12. The prevalence effect is determined by past experience, not future prospects

Jonas Sin-Heng Lau, Liqiang Huang

Vision Research (2010-07) <https://doi.org/bj9p7j>

DOI: [10.1016/j.visres.2010.04.020](https://doi.org/10.1016/j.visres.2010.04.020) · PMID: [20438744](https://pubmed.ncbi.nlm.nih.gov/20438744/)

13. Homo economicus in visual search

V. Navalpakkam, C. Koch, P. Perona

Journal of Vision (2009-01-01) <https://doi.org/dnjhzz>

DOI: [10.1167/9.1.31](https://doi.org/10.1167/9.1.31) · PMID: [19271901](https://pubmed.ncbi.nlm.nih.gov/19271901/)

14. False feedback increases detection of low-prevalence targets in visual search

Jeremy Schwark, Joshua Sandry, Justin MacDonald, Igor Dolgov

Attention, Perception, & Psychophysics (2012-08-03) <https://doi.org/f4gqwh>

DOI: [10.3758/s13414-012-0354-4](https://doi.org/10.3758/s13414-012-0354-4) · PMID: [22864899](https://pubmed.ncbi.nlm.nih.gov/22864899/)

15. Prevalence effects in newly trained airport checkpoint screeners: Trained observers miss rare targets, too

J. M. Wolfe, D. N. Brunelli, J. Rubinstein, T. S. Horowitz

Journal of Vision (2013-12-02) <https://doi.org/gfkzhq>

DOI: [10.1167/13.3.33](https://doi.org/10.1167/13.3.33) · PMID: [24297778](https://pubmed.ncbi.nlm.nih.gov/24297778/) · PMCID: [PMC3848386](https://pubmed.ncbi.nlm.nih.gov/PMC3848386/)

-
1. We use the term “medical image perception” to cover a wide range of imaging techniques from images of individual cells used in pathology to whole-body radiological images, from glass slides to 3D computer renderings. In addition to detection and diagnosis, medical imaging can also be used to guide surgical interventions. We also extend to term beyond artificially created images; a dermatologist examining your skin, for example, is engaged in a medical image perception task. ↵
 2. Here we will use the word “clinicians” to refer to anyone tasked with reading a medical image. The typical clinician in this sense will be a radiologist or pathologist, but the term should be understood to include many other specialties such as nuclear medicine or dermatology, as well as non-specialists who read images as part of their practice. ↵