



The invisible breast cancer: Experience does not protect against inattentional blindness to clinically relevant findings in radiology

Lauren Williams¹ · Ann Carrigan² · William Auffermann³ · Megan Mills³ · Anina Rich⁴ · Joann Elmore⁵ · Trafton Drew¹

Accepted: 4 October 2020
© The Psychonomic Society, Inc. 2020

Abstract

Retrospectively obvious events are frequently missed when attention is engaged in another task—a phenomenon known as inattentional blindness. Although the task characteristics that predict inattentional blindness rates are relatively well understood, the observer characteristics that predict inattentional blindness rates are largely unknown. Previously, expert radiologists showed a surprising rate of inattentional blindness to a gorilla photoshopped into a CT scan during lung-cancer screening. However, inattentional blindness rates were higher for a group of naïve observers performing the same task, suggesting that perceptual expertise may provide protection against inattentional blindness. Here, we tested whether expertise in radiology predicts inattentional blindness rates for unexpected abnormalities that were clinically relevant. Fifty radiologists evaluated CT scans for lung cancer. The final case contained a large (9.1 cm) breast mass and lymphadenopathy. When their attention was focused on searching for lung nodules, 66% of radiologists did not detect breast cancer and 30% did not detect lymphadenopathy. In contrast, only 3% and 10% of radiologists ($N = 30$), respectively, missed these abnormalities in a follow-up study when searching for a broader range of abnormalities. Neither experience, primary task performance, nor search behavior predicted which radiologists missed the unexpected abnormalities. These findings suggest perceptual expertise does not protect against inattentional blindness, even for unexpected stimuli that are within the domain of expertise.

Keywords Attention · Visual perception · Visual search · Inattentional blindness

Despite our intuition that visual perception provides a faithful representation of the world, we often miss salient events in our environment when we are focused on something else. This phenomenon is known as inattentional blindness—the failure to notice an unexpected but fully visible stimulus when attention is engaged in another task (Mack & Rock, 1998; Neisser & Becklen, 1975). The most famous example of inattentional blindness is the “invisible gorilla” (Simons & Chabris, 1999).

When observers were asked to watch a ball game and count the number of ball passes between teammates, ~50% of observers missed a person in a gorilla-costume walking through the middle of the game and beating its chest. In homage to this work, Drew, Võ, and Wolfe (2013) inserted a matchbook-sized image of a gorilla into a chest computed tomography (CT) scan and asked radiologists to look for lung-cancer nodules. Under these conditions, 83% of radiologists did not notice the gorilla, demonstrating that sustained inattentional blindness can occur at high rates in expert observers performing a task for which they are highly trained.

Although there are several known situational factors that affect inattentional blindness rates, whether some individuals are more susceptible to inattentional blindness than others is a substantial gap in the literature (Simons & Jensen, 2009). Many likely observer characteristics, such as working memory capacity and primary task performance, fail to predict which individuals will experience inattentional blindness (Bredemeier & Simons, 2012; Kreitz, Furley, Memmert, & Simons, 2015; Simons & Jensen, 2009). Consequently, some

✉ Lauren Williams
laurenhayleywilliams@gmail.com

¹ Psychology, University of Utah, Salt Lake City, UT, USA

² Psychology, Macquarie University, Macquarie Park, Australia

³ School of Medicine, Department of Radiology and Imaging Sciences, University of Utah, Salt Lake City, UT, USA

⁴ Cognitive Science, Macquarie University, Macquarie Park, Australia

⁵ David Geffen School of Medicine, UCLA, Los Angeles, CA, USA

researchers have proposed that inattention blindness is largely a stochastic process rather than a stable individual trait (Kreitz et al., 2015; Simons & Jensen, 2009). However, other studies suggest that expertise on the primary task may provide some protection against inattention blindness. Basketball experience was associated with reduced inattention blindness rates when viewing a ball-playing game (Memmert, 2006). Similarly, our prior work found that expert radiologists exhibited a lower rate of inattention blindness than naïve undergraduate observers performing the same task (Drew, Vö, & Wolfe, 2013).

Experience may be associated with lower inattention blindness rates because experts do not need to devote as many cognitive resources to the primary task (Fougnie & Marois, 2007). Specifically, inattention blindness rates are predicted by relative task difficulty (i.e., effort) rather than an individual's performance on the primary task per se (Simons & Jensen, 2009). Alternatively, experts may benefit from enhanced perceptual abilities or superior search strategies for tasks within their domain of expertise (Memmert, 2006). For example, enhanced global processing ability in expert radiologists allows them to rapidly detect abnormalities, which might make them more likely to detect the unexpected abnormalities (Carrigan, Wardle, & Rich, 2018; Evans, Georgian-Smith, Tambouret, Birdwell, & Wolfe, 2013; Drew, Evans, Vö, Jacobson, & Wolfe, 2013). In previous research, the beneficial effects of experience have only been demonstrated in groups of experts relative to naïve observers (Drew et al., 2013; Memmert, 2006). Thus, it remains unclear whether the observed individual differences are related to domain-specific perceptual abilities acquired through experience (Memmert, 2006) or a higher cognitive load in novices learning a completely new task (Fougnie & Marois, 2007). To address this question, it is necessary to determine whether experience predicts inattention blindness within a group of observers that are familiar with the primary task. Furthermore, if expertise does protect against inattention blindness, it is important to differentiate between the various factors that might be driving these effects (e.g., search strategy).

Our prior work suggests that highly unusual abnormalities (e.g., gorillas in a CT scan) may be frequently missed by radiologists if they are unexpected (Drew et al., 2013). However, because radiologists never need to identify images of gorillas in the lung in clinical practice, it may be that clinical experience protects against inattention blindness to unexpected stimuli that relate to the primary task and to the observer's experience. Inattention blindness rates are lower for unexpected objects that are more similar to the observer's attentional set (Most et al., 2001). This modulation occurs for objects with shared physical features (Most et al., 2001), as well as dissimilar objects from the same semantic category (Koivisto & Revonsuo, 2007; Most, 2013). Therefore, radiologists might be less likely to miss clinically relevant abnormalities than unexpected stimuli that are unusual within the

context of the task. Furthermore, experts are thought to have greater attentional capacity for tasks within their area of expertise, rather than enhanced domain-general abilities (Memmert, Simons, & Grimme, 2009). The perceptual expertise of radiologists may be finely tuned to clinically relevant features, which might make them less likely to miss unexpected stimuli relevant to their area of expertise. In addition, an important characteristic of inattention blindness is the unexpected nature of the secondary stimulus (White & Davies, 2008; cf. Ward & Scholl, 2015). However, radiologists are incentivized to identify unexpected abnormalities that are clinically relevant, such as tumors, which could endanger their patient and lead to costly litigation. This makes inattention blindness in radiology a unique perspective for studying the classic inattention blindness phenomenon.

The purpose of this research was to provide a stronger test of the proposal that primary task experience provides protection against inattention blindness. In Study 1, we recruited a large sample of radiologists to examine the effect of perceptual expertise on inattention blindness for stimuli that fit within the experience set of the participants and the task at hand: detection of a breast mass and lymphadenopathy during lung-cancer screening. In addition, we performed an exploratory analysis to determine whether radiologists who detected the unexpected abnormalities engaged in different strategies, such as searching the images more thoroughly, which would provide more opportunities for abnormality detection. In Study 2, we sought to establish baseline detection rates for the secondary abnormalities when radiologists focused their attention on a broader set of abnormalities.

Study 1

Method

Participants and overview

Fifty-six practicing radiologists and residents in radiology training programs were recruited from the Radiological Society for North America meeting; a hospital in Salt Lake City, UT; and a hospital in Sydney, Australia. Six radiologists were excluded from the study due to equipment failure.

To improve the reliability and reproducibility of the results, the sample size, exclusion criteria, and primary predictions were preregistered prior to data collection as part of a larger study (<https://osf.io/aqkbd>) (Adibi, Sin, & Sadatsafavi, 2019). The preregistered sample size was based on previous research on the effects of experience on inattention blindness in radiology (Drew et al., 2013). In addition to our preregistered analyses, we calculated Bayes factors to assist in the interpretation of null effects (Jeffreys, 1998). The study was approved by the University of Utah Institutional Review Board and the

Macquarie University Human Research Ethics Committee. All participants provided informed consent.

Materials

Observers evaluated seven axial chest CT scans (one practice and six experimental) in a typical lung window level. Three experimental cases were normal and three contained lung nodules. The final case contained a large breast mass and lymphadenopathy in addition to lung-cancer (see Fig. 1). The mass was 9.1cm in maximum diameter and was visible on 17/66 slices. At its maximum diameter, the breast mass was approximately 30 times the size of the smallest lung nodules. This case, obtained from clinical practice of author W.A., was chosen because the abnormalities were deemed clearly visible using lung window level settings. Our goal was to use a case with abnormalities that would be easily detected and unambiguous if the radiologist were explicitly looking for them. The remaining images and nodule locations were obtained from the Lung Image Database Consortium (LIDC; Armato et al., 2011). The practice case had eight lung nodules, and the abnormal experimental cases contained 14.33 nodules on average (9, 11, & 23 per case). The final, critical case had 11 nodules. For the cases obtained from the LIDC, ground truth was established by four thoracic radiologists that marked nodule locations prior to reviewing the anonymized marks of the other three radiologists and

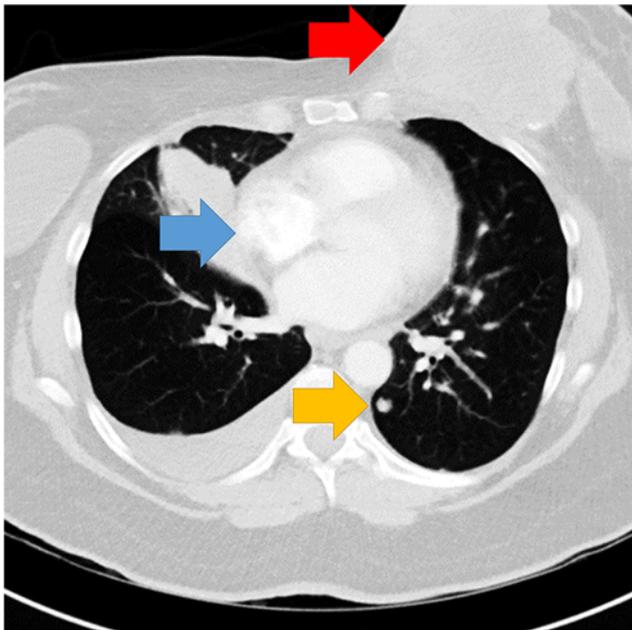


Fig. 1 Single-slice of the chest CT showing the abnormalities. Arrows indicate the location of the breast mass (red arrow), lymphadenopathy (blue arrow), and a lung nodule (yellow arrow). Arrows not present in experimental display

rendering a final decision (Armato et al., 2011). For the final case, author W.A. marked the nodule locations.

Procedure

Observers were asked to perform a lung nodule screening task and were instructed “to mark all nodules ≥ 3 mm” by clicking on them with the mouse. The radiologists were not given additional instructions regarding any secondary abnormalities that might be present in the image. Observers were able to freely scroll through depth using the scroll wheel. Responses were considered correct if a click was located within 30 pixels and two slices of the nodule’s center of mass. Cases were presented without any additional patient data, and the radiologists were not asked to provide any additional diagnosis beyond marking the location of nodules until after the final case. The practice case and the critical case were always presented in the same order (first and last), and the order of the other cases was randomized for each participant. Following the final case, observers answered a series of yes/no questions on the computer to assess for possible inattentive blindness:

- 1) Did the final case seem any different than any of the other trials?
- 2) Did you notice any other medically relevant findings on the final case?
- 3) Did the final case show signs of breast cancer?
- 4) Did the final case show signs of lymphadenopathy?

Questions 1 and 2 were presented in the set order, whereas the order of Questions 3 and 4 was randomized across participants.

The images subtended $\sim 15.15^\circ$ of visual angle and were displayed on a 17-in. laptop (resolution: $1,600 \times 900$, refresh rate: 60Hz) using MATLAB with Psychtoolbox (Brainard, 1997). Participants were situated on a chin rest ~ 89 cm from the monitor. Eye movements were recorded at 1000 Hz using an EyeLink 1000 Plus, and the current slice was coregistered with each sample.

Analysis plan

We divided radiologists into groups based on whether they reported the unexpected abnormalities (i.e., a yes response to either Question 3 or 4). We compared groups using a two-tailed, between-participants *t* test. This analysis deviated from the preregistered plan to use linear regression for this data set because inattentive blindness is a binary measure. For each comparison, we additionally computed Bayes factors (BF). A $BF_{01} > 3$ was considered at least moderate evidence for the null relative to the alternative hypothesis, and a $BF_{01} < .33$ was considered at least moderate evidence for the alternative relative to the null hypothesis. Bayes factors between .33 and

3 indicate there is insufficient evidence to merit a strong conclusion (Jeffreys, 1998). Finally, we computed a logistic regression to determine whether experience predicted inattentive blindness while controlling for other variables (e.g., age). For those analyses, we reported the odds ratio and the p value from the Wald Test.

Results

The final sample of 50 radiologists included 25 radiology residents in training, one fellow in radiology, and 24 attending/practicing radiologists. Of these radiologists, 20 (40%) reported chest imaging as a specialty. The average radiologist in our study completed medical school 12.35 years ago ($SD = 12.58$, range: .50–42) and reported reading 41 chest CT scans in a typical week ($SD = 52$, range: 0–250). The average age was 41 years ($SD = 12.88$). Trainees (residents and fellows) had 4.09 years of experience ($SD = 5.88$) and reported reading 46.54 scans each week ($SD = 54.70$), whereas attendings had 21.29 years of experience ($SD = 11.79$) and reported reading 34.75 chest CT scans each week ($SD = 49.33$). On average, radiologists detected 58.38% ($SD = 18.5%$, range: 14.01%–83.25%) of the lung nodules.

Detection of incidental breast cancer

When searching for lung cancer, 33/50 (66%) radiologists did not identify a large, clinically relevant breast mass (see Fig. 2a).¹ These radiologists did not differ from radiologists who successfully detected the breast cancer in years of experience, $t(48) = .18, p = .86, BF_{01} = 3.35$, or lung-nodule detection rate, $t(48) = .35, p = .73, BF_{01} = 3.22$, and there was weak evidence for no difference in the number of chest CTs read per week, $t(47) = 1.46, p = .15, BF_{01} = 1.45$ (see Fig. a–c, Table 1). In a logistic regression, neither years of experience (odds ratio = .89, $p = .24$), age (odds ratio = 1.12, $p = .25$), nor the number of chest CTs read per week (odds ratio = .99, $p = .09$) predicted breast cancer detection. Chest imaging experts were no more likely to detect the breast cancer mass than nonchest experts ($\chi^2 = .53, p = .46$). In exploratory analyses, we found some evidence that radiologists who detected the mass spent more time looking at the abnormality than those who did not, $t(48) = 2.23, p = .03, BF_{01} = .48$, but this did not reach our threshold for sufficient evidence ($BF_{01} < 0.33$; see Table 1).

¹ Of the 17 radiologists that reported the breast mass, one responded “no” to Question 1 (“Did the final case seem any different than any of the other trials?”) only, one responded “no” to Question 2 (“Did you notice any other medically relevant findings on the final case?”) only, and one responded “no” to both questions. Of the 35 radiologists that reported lymphadenopathy, three responded “no” to Question 1 only, two responded “no” to Question 2 only, and one responded “no” to both questions. Although it is impossible to know for certain based on a yes or no response to these questions, this pattern of responses may reflect a demand characteristic for a small number of participants when asked directly about specific abnormalities.

There does not appear to be a difference between the radiologists who missed the mass and those that detected it in case completion time, $t(48) = 1.44, p = .16, BF_{01} = 1.47$, or the percentage of the image examined, $t(48) = 1.12, p = .27, BF_{01} = 2.04$ (see Table 1), but the Bayes factors indicate more evidence is needed.

Detection of incidental lymphadenopathy

Thirty-percent (15/50) of radiologists reported no signs of lymphadenopathy (see Fig. 2a). The breast mass was missed more often than the lymphadenopathy (Fisher’s exact test, $p = .001$). These radiologists did not differ from radiologists who reported the lymphadenopathy in years of experience, $t(48) = .36, p = .72, BF_{01} = 3.14$, and there was insufficient evidence regarding the number of chest CTs read per week, $t(47) = 1.42, p = .16, BF_{01} = 1.47$, and lung-nodule detection-rate, $t(48) = 1.06, p = .29, BF_{01} = 2.11$, (see Fig. 3d–f, Table 2). In a logistic regression, neither years of experience (odds ratio = .92, $p = .34$), age (odds ratio = 1.08, $p = .38$), nor the number of chest CTs read per week (odds ratio = .99, $p = .14$) predicted breast cancer detection while controlling for the other variables. Chest imaging experts were no more likely to detect the lymphadenopathy than were nonchest experts ($\chi^2 = 1.59, p = .21$). Observers who detected the abnormality spent less time searching the image than those who did not, $t(48) = 2.025, p = .0485, BF_{01} = .66$, but this effect did not reach our level for substantial evidence ($BF_{01} < 0.33$). We had insufficient evidence to determine whether there were differences in image coverage, $t(48) = 1.43, p = .16, BF_{01} = 1.47$ (see Table 2).

Study 2

In Study 1, the critical case was chosen because the secondary abnormalities seemed clearly visible. However, it is possible these abnormalities would have been missed at high rates even if the observers were not asked to focus on a narrow clinical task of evaluating only lung nodules. To establish whether these effects are attentional in nature, we performed a second study to determine the baseline detection rates for the critical abnormalities when attention was focused more broadly.

Method

Participants

Thirty-three radiologists were recruited from hospitals in Salt Lake City, UT, and Sydney, Australia, by emailing radiologists who had participated in previous studies, advertising to radiology-oriented groups on social media, and through word-of-mouth. Radiologists that participated in Study 1 were not recruited to participate in Study 2. Three radiologists were

Table 1 Descriptive statistics broken down by whether breast cancer was or was not detected

Breast cancer detection		Detected abnormality	Mean	SD	95% CI mean difference	<i>p</i> value	BF ₀₁
Years of experience	Yes		12.79	13.34	[-8.30, 6.96]	.86	3.35
	No		12.12	12.38			
Chest CTs per week	Yes		55.41	71.53	[-53.43, 8.57]	.15	1.45
	No		32.98	36.83			
% lung nodules found	Yes		59.67%	18.25%	[13.15%, 9.25%]	.73	3.22
	No		57.71%	18.86%			
Time spent looking at abnormality	Yes		2,320 ms	4,272 ms	[-3,423; -176.6 ms]	.03	.48
	No		519.8 ms	1,359 ms			
Case completion time	Yes		145.3 s	58.23 s	[-57.10 s, 9.38 s]	.16	1.47
	No		121.5 s	53.89 s			
% of image searched	Yes		49.36%	19.05%	[-14.02%, 3.99%]	.27	2.04
	No		44.34%	12.49%			

excluded from the study: two did not meet the inclusion criteria (at least a first-year radiology resident) and one did not scroll through the CT scans before providing a response.

The study's sample size, analysis plan, and exclusion criteria were preregistered (<https://osf.io/73en6>). The study was approved by the University of Utah Institutional Review Board, and all participants provided informed consent.

Materials and procedure

The study was conducted online using a nonmobile device of the radiologist's choosing. Observers evaluated two axial chest CT scans (one normal case and one abnormal case) in a typical lung window level. The abnormal case was the critical case from Study 1, which allowed us to establish a baseline detection rate for each of the critical abnormalities. The order of the normal and abnormal cases was counterbalanced. For each case, observers were asked, "Which abnormalities are present in this case (check all that apply)?: lung nodules, a breast mass, lymphadenopathy, pneumonia, pulmonary

embolism, or no major findings. The order of the response options was randomized for each participant. Participants could scroll through the image using a scroll wheel or the arrows on their keyboard.

Results

The final Study 2 sample of 30 radiologists included five radiology residents in training, five radiology fellows, and 20 attending/practicing radiologists. Five (~17%) of these radiologists reported chest imaging as a specialty. On average, the radiologists had 10 ($SD = 9.90$, range: 2, 43) years of experience and reported reading 47.6 ($SD = 68.17$, range: 0, 300) chest CT scans per week.

When asked about the critical abnormalities directly, only 1/30 (3%) radiologists missed the breast mass and 3/30 (10%) missed the lymphadenopathy (see Fig. 2b). Fisher's exact tests revealed the breast mass was detected more often in Study 2 than in Study 1, $p < .001$, $BF_{01} = 1.57e-6$. The difference in detection rate for the lymphadenopathy between the two studies was only

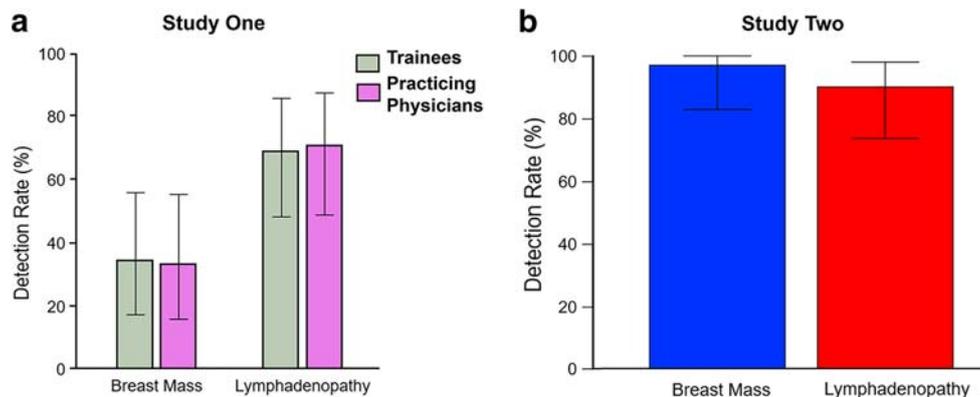


Fig. 2 **a** Detection rate for breast mass and lymphadenopathy by level of training in Study 1. **b** Detection rate for breast mass and lymphadenopathy in Study 2. Error bars represent the 95% confidence intervals

Table 2 Descriptive statistics broken down by whether lymphadenopathy was or was not detected

Lymphadenopathy detection		Detected abnormality	Mean	SD	95% CI mean difference	<i>p</i> value	BF ₀₁
Years of experience	Yes	Yes	12.77	12.63	[-9.28, 6.47]	.72	3.14
	No	No	11.37	12.86			
Chest CTs per week	Yes	Yes	47.36	58.87	[-55.77, 9.63]	.16	1.47
	No	No	24.29	22.09			
% lung nodules found	Yes	Yes	60.19%	18.85%	[-17.52%, 5.41%]	.29	2.11
	No	No	54.14%	17.52%			
Case completion time	Yes	Yes	139.8 s	60.03 s	[-67.62 s, .23 s]	.049	.66
	No	No	105.9 s	36.83 s			
% of image searched	Yes	Yes	48.02%	16.40%	[-15.80%, 2.67%]	.16	1.47
	No	No	41.45%	10.31%			

marginally significant, $p = .053$, $BF_{01} = .22$. However, the Bayes factor met our criteria for sufficient evidence in favor of a higher detection-rate in Study 2 than Study 1. Observers also performed well on the lung-cancer detection task: Only 1/30 (3%) radiologists missed the lung nodules. In all, 29/30 (97%) radiologists reported there were no signs of pulmonary embolism, and 16/30 (53%) radiologists reported there were no signs of pneumonia. No radiologists reported a breast mass or lymphadenopathy in the normal case. McNemar's test revealed these reports were significantly lower than the abnormal case, both p values $< .001$, which demonstrates radiologists only reported the critical abnormalities when they were actually present.

Discussion

Inattention blindness in radiology is not limited to missing an image of a gorilla embedded in a CT scan (Drew et al., 2013). Two-thirds of radiologists did not detect a clinically relevant 9.1-cm breast mass when their attention was engaged in an attentionally demanding lung-nodule-detection task. In contrast, only 3% of radiologists missed the mass when directly asked about these abnormalities. These findings demonstrate that inattention blindness is not limited to completely unexpected stimuli (e.g., a gorilla in a CT scan): The unexpected abnormalities in our study were clinically relevant and within the observer's skill set. Critically, these findings suggest perceptual expertise does not provide protection against

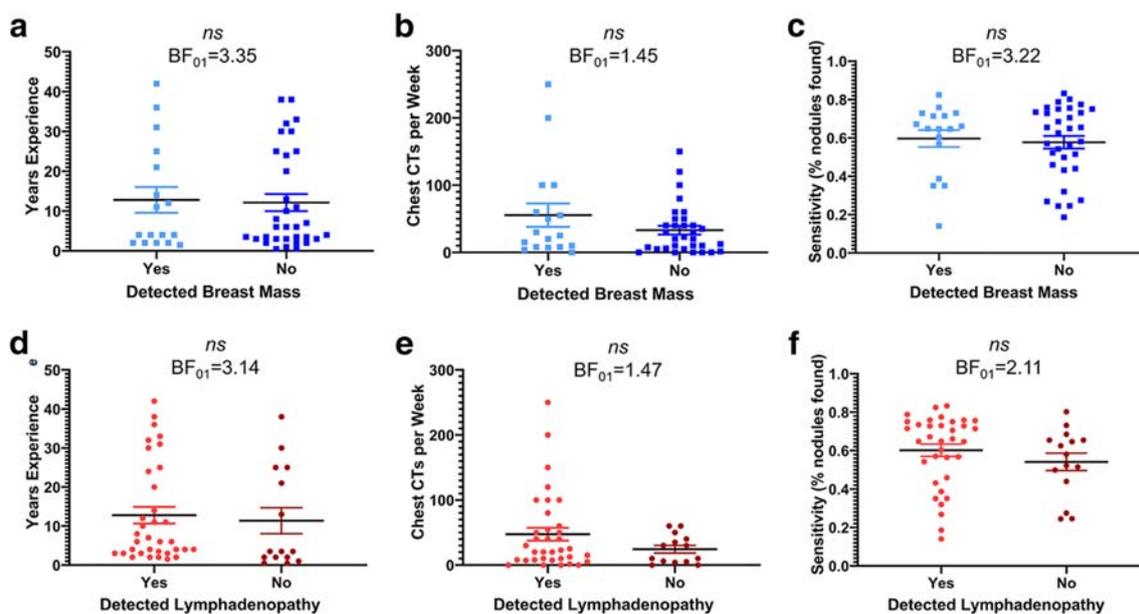


Fig. 3 Clinical experience and sensitivity in detection of lung nodules shown by detection of the incidental breast mass (panels a, b, c) and detection of incidental lymphadenopathy (panels d, e, f). The solid line

indicates the mean value, dots represent the individual data points, and error bars represent standard error of the mean. A $BF_{01} > 3$ indicates there is sufficient evidence for the null relative to the alternative hypothesis

inattentional blindness: radiologists who missed the unexpected abnormalities did not differ in years of experience or overall task performance. Although task familiarity appears to modulate inattentional blindness rates to an extent (Drew et al., 2013; Memmert, 2006), this finding may be due to the relative difficulty of performing the task rather than domain-specific perceptual expertise (Simons & Jensen, 2009). Our results indicate inattentional blindness may be an inherent limitation of the cognitive system, and expertise does not protect against this phenomenon.

Inattentional blindness may provide insight on how to address a growing concern in radiology: missed but clinically significant incidental findings. Incidental findings are abnormalities in medical images that are unrelated to the patient's symptoms or present in an asymptomatic patient (O'Sullivan, Muntinga, Grigg, & Ioannidis, 2018). For example, a medical image taken to determine if a patient has a broken bone might incidentally reveal signs suggestive of cancer. Regardless of the reason the image was obtained, radiologists are expected to report all clinically significant findings. Missed or incorrect diagnoses impact the quality of patient care and are the leading cause of medical malpractice litigation (Tehrani et al., 2013). A recent meta-analysis found that incidental findings are present on ~1/4 diagnostic images (Lumbreras, Donat, & Hernández-Aguado, 2010) but are often undocumented, unreported to patients, and not followed-up for additional testing (Munk, Peitzman, Hostler, & Wolfson, 2010). With medical errors estimated to be the third most common cause of death in the United States (Makary & Daniel, 2016), it is vital to better understand the cognitive mechanisms underlying this source of diagnostic error.

Our data suggest the likelihood of experiencing inattentional blindness is strongly influenced by whether the unexpected abnormality occurs in a task-relevant location. The final case was chosen because both abnormalities were clearly visible in the viewing window used for lung-cancer screening. However, detection-rate for the large breast mass was significantly lower than the lymphadenopathy in Study 1, which may be because the mass was located outside of the chest wall, whereas the lymphadenopathy was located within the chest wall. Clinicians may adopt the strategy of attending less to areas outside the primary organs of interest when engaged in a difficult task. This pattern of data is consistent with previous findings: Unexpected objects are less likely to be noticed when they appear further away from the focus of attention (Most, Simons, Scholl, & Chabris, 2000; Wood & Simons, 2019). Similarly, abnormalities outside of the task-relevant location, such as lung cancer on musculoskeletal radiographs, are often retrospectively visible on images mistakenly classified as normal in clinical practice (Hawkesford & Kalogrianitis, 2015; Kim & Mansfield, 2014). This retrospective identification may be more akin to Study 2, when

participants were specifically asked about different abnormalities. Under these conditions, radiologists' detection rates for the breast mass increased from 33% to 97%, and there was a similar effect for lymphadenopathy: detection rates were lower during lung-cancer screening (70%) than a more general task (90%).

This research complements previous work on the common mechanisms of missed targets in visual search. Previous studies show a tendency to miss a second target after the detection of the first target (Berbaum et al., 1990; Tuddenham, 1962). Initially attributed to a premature termination of search, these "subsequent search misses" (Cain, Adamo, & Mitroff, 2013) remain a stubborn source of error (Cheng & Rich, 2018). In our study, radiologists who did and did not detect the abnormalities had similar task-completion time and image coverage, which suggests these misses also cannot be attributed to premature case closure. However, there are important differences between our results and subsequent search misses (Wolfe, Soce, & Schill, 2017). Subsequent search misses are the failure to detect a known target, typically of the same type of target as the first (e.g., two lung nodules). Here, radiologists missed unspecified abnormalities that were not the focus of the search. This phenomenon meets the criteria for inattentional blindness established by Rock, Linnett, Grant, and Mack (1992): (1) The majority of observers failed to notice the abnormalities, (2) the abnormalities were fully visible, (3) observers were able to identify the abnormalities when asked about them directly (Study 2), and (4) the abnormalities were unexpected, and they were missed more often when attention was focused on another task (Study 1) than when it was not (Study 2). More research is needed to determine how the mechanisms behind these two types of misses are related.

There are some limitations to this study. The radiologists viewed these images in a laboratory setting where they knew their decisions would not affect any patients, and our inattentional blindness paradigm differs in important ways from clinical practice. Radiologists were engaged in the same clinical task for multiple consecutive trials before the critical case. Although common in screening tasks, it remains to be seen if the observed effects are influenced by prior task history. Our results are based on a single case that was only viewed in the lung window, and more work is necessary to determine whether these findings generalize. However, we deliberately chose a case with a large, clearly visible mass, and we anticipate detection rate would be even lower for more subtle abnormalities. In future work, it will be beneficial to replicate these results in a clinical setting with a broader set of stimuli. Finally, Study 2 was conducted online, which may not match the viewing conditions and observer state of mind of Study 1. For example, the monitor specifications (i.e., resolution and brightness) may have differed between

studies. Nonetheless, almost every radiologist detected the critical abnormalities, clearly demonstrating these abnormalities are visible and easily detectable when attention is focused more broadly.

Our research demonstrates the powerful influence attention can have on medical image interpretation. Even if an image shows clear signs of an abnormality, the clinician may fail to perceive the finding if it is inconsistent with the reason the exam was obtained. This result is increasingly important considering the recent push for radiologists to focus their attention on specific tasks to reduce the number of nonclinically relevant incidental findings reported in radiology (Barry, 2014). This research has important implications for considering how to best present relevant information (e.g., patient history) to the clinician and how to evaluate whether “standard of care” has been violated in medical malpractice cases. If most radiologists do not detect a large, clinically relevant cancer while searching for a different type of abnormality, we should acknowledge that the concept of “standard of care” may need to be expanded to include the exam indication and the instructions provided. However, these results should not be interpreted as evidence that it would be better not to provide patient information. Our results suggest the information given to the radiologist may have a strong influence on which abnormalities they will ultimately detect. Therefore, relevant patient information should be helpful to radiologists when the abnormalities are consistent with the exam indication. In fact, a large body of research demonstrates that clinical history improves sensitivity for relevant abnormalities (Loy & Irwig, 2004). Rather than providing no information, it may be helpful to focus on interventions, such as checklists (Thomassen, Storesund, Søfteland, & Brattebø, 2014), that help clinicians focus attention on one task at a time rather than trying to divide attention between multiple competing goals.

Our findings provide a unique perspective on classic inattention blindness, demonstrating it occurs when the unexpected stimulus is by no means unusual within the context, and that perceptual expertise cannot prevent it. In sum, this is a vivid illustration of Mack and Rock’s (1998) assertion that even with trained experts viewing an image within their area of expertise, “there is no perception without attention.”

Open practices statement

Both studies were preregistered on the Open Science Framework (Study 1: <https://osf.io/aqkbd>, Study 2: <https://osf.io/73en6>). The data have not been made available on a permanent third-party archive; requests for the data can be sent via email to the lead author. The experiment script and expertise questionnaire can be found at https://osf.io/je8tm/?view_only=927401fb4b28449fb31bd665fcb4acb5

Acknowledgements We would like to thank the radiologists that participated in our studies, the National Cancer Institute’s Perception Lab at the annual meeting of the Radiological Society of North America, Daniel J. Williams for creating the online CT viewer in Study 2, and two anonymous reviewers for their helpful comments during the review process.

References

- Adibi, A., Sin, D., & Sadatsafavi, M. (2019). Lowering the p value threshold. *JAMA*, *321*(15), 1532–1533. doi:<https://doi.org/10.1001/jama.2019.0566>
- Armato, S. G., McLennan, G., Bidaut, L., McNitt-Gray, M. F., Meyer, C. R., Reeves, A. P., ... Croft, B. Y. (2011). The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed reference database of lung nodules on CT scans. *Medical Physics*, *38*(2), 915–931. doi:<https://doi.org/10.1118/1.3528204>
- Barry, M. J. (2014). Incidentaloma Fatigue. *JAMA Internal Medicine*, *174*(6), 851–852. doi:<https://doi.org/10.1001/jamainternmed.2014.672>
- Berbaum, K. S., Franken, J. E., Dorfman, D. D., Rooholamini, S. A., Kathol, M. H., & Barloon, T. J. (1990). Satisfaction of search in diagnostic radiology. *Investigative Radiology*, *25*. doi:<https://doi.org/10.1097/00004424-199002000-00006>
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436.
- Bredemeier, K., & Simons, D. J. (2012). Working memory and inattention blindness. *Psychonomic Bulletin & Review*, *19*(2), 239–244.
- Cain, M. S., Adamo, S. H., & Mitroff, S. R. (2013). A taxonomy of errors in multiple-target visual search. *Visual Cognition*, *21*(7), 899–921.
- Carrigan, A. J., Wardle, S. G., & Rich, A. N. (2018). Finding cancer in mammograms: If you know it’s there, do you know where? *Cognitive Research: Principles and Implications*, *3*.
- Cheng, P. X., & Rich, A. N. (2018). More is better: Relative prevalence of multiple targets affects search accuracy. *Journal of Vision*, *18*(4), 2–2. doi:<https://doi.org/10.1167/18.4.2>
- Drew, T., Evans, K., Vö, M. L. H., Jacobson, F. L., & Wolfe, J. M. (2013). Informatics in radiology: What can you see in a single glance and how might this guide visual search in medical images? *Radiographics*, *33*. <https://doi.org/10.1148/rg.331125023>
- Drew, T., Vö, M. L. H., & Wolfe, J. M. (2013). The invisible gorilla strikes again: Sustained inattention blindness in expert observers. *Psychological Science*, *24*. <https://doi.org/10.1177/0956797613479386>
- Evans, K. K., Georgian-Smith, D., Tambouret, R., Birdwell, R. L., & Wolfe, J. M. (2013). The gist of the abnormal: Above-chance medical decision making in the blink of an eye. *Psychonomic Bulletin & Review*, *20*. doi:<https://doi.org/10.3758/s13423-013-0459-3>
- Fougnie, D., & Marois, R. (2007). Executive working memory load induces inattention blindness. *Psychonomic Bulletin & Review*, *14*(1), 142–147.
- Hawkesford, M. P. H., & Kalogrianitis, S. (2015). Delayed diagnosis of lung cancer after missed vertebral metastasis on CT. *BJR Case Reports*, *1*(1). doi:<https://doi.org/10.1259/bjrcr.20140001>
- Jeffreys, H. (1998). *The theory of probability*. Oxford, UK: Oxford University Press.
- Kim, Y. W., & Mansfield, L. T. (2014). Fool me twice: Delayed diagnoses in radiology with emphasis on perpetuated errors. *American Journal of Roentgenology*, *202*(3), 465–470. doi:<https://doi.org/10.2214/AJR.13.11493>
- Koivisto, M., & Revonsuo, A. (2007). How meaning shapes seeing. *Psychological Science*, *18*(10), 845–849. doi:<https://doi.org/10.1111/j.1467-9280.2007.01989.x>

- Kreitz, C., Furley, P., Memmert, D., & Simons, D. J. (2015). Inattentional Blindness and Individual Differences in Cognitive Abilities. *PLOS ONE*, *10*(8), e0134675. doi:<https://doi.org/10.1371/journal.pone.0134675>
- Loy, C. T., & Irwig, L. (2004). Accuracy of diagnostic tests read with and without clinical information: A systematic review. *Jama*, *292*(13), 1602–1609.
- Lumbreras, B., Donat, L., & Hernández-Aguado, I. (2010). Incidental findings in imaging diagnostic tests: A systematic review. *The British Journal of Radiology*, *83*(988), 276–289. doi:<https://doi.org/10.1259/bjr/98067945>
- Mack, A., & Rock I. (1998). *Inattentional blindness*. MIT Press.
- Makary, M. A., & Daniel, M. (2016). Medical error—The third leading cause of death in the US. *BMJ*, *353*, i2139. doi:<https://doi.org/10.1136/bmj.i2139>
- Memmert, D. (2006). The effects of eye movements, age, and expertise on inattentional blindness. *Consciousness and Cognition*, *15*(3), 620–627.
- Memmert, D., Simons, D. J., & Grimme, T. (2009). The relationship between visual attention and expertise in sports. *Psychology of Sport and Exercise*, *10*(1), 146–151.
- Most, S. B. (2013). Setting sights higher: Category-level attentional set modulates sustained inattentional blindness. *Psychological Research*, *77*(2), 139–146. doi:<https://doi.org/10.1007/s00426-011-0379-7>
- Most, S. B., Simons, D. J., Scholl, B. J., & Chabris, C. F. (2000). Sustained inattentional blindness. *Psyche*, *6*, 14.
- Most, S. B., Simons, D. J., Scholl, B. J., Jimenez, R., Clifford, E., & Chabris, C. F. (2001). How not to be Seen: The Contribution of Similarity and Selective Ignoring to Sustained Inattentional Blindness. *Psychological Science*, *12*(1), 9–17. doi:<https://doi.org/10.1111/1467-9280.00303>
- Munk, M.-D., Peitzman, A. B., Hostler, D. P., & Wolfson, A. B. (2010). Frequency and follow-up of incidental findings on trauma computed tomography scans: Experience at a level one trauma center. *The Journal of Emergency Medicine*, *38*(3), 346–350. doi:<https://doi.org/10.1016/j.jemermed.2008.01.021>
- Neisser, U., & Becklen, R. (1975). Selective looking: Attending to visually specified events. *Cognitive Psychology*, *7*(4), 480–494.
- O’Sullivan, J. W., Muntinga, T., Grigg, S., & Ioannidis, J. P. A. (2018). Prevalence and outcomes of incidental imaging findings: Umbrella review. *BMJ*, *361*, k2387. doi:<https://doi.org/10.1136/bmj.k2387>
- Rock, I., Linnett, C.M., Grant, P., & Mack, A. (1992) Perception without attention: Results of a new method. *Cognitive Psychology*, *24*(4), 502–534.
- Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception*, *28*(9), 1059–1074. doi:<https://doi.org/10.1068/p281059>
- Simons, D. J., & Jensen, M. S. (2009). The effects of individual differences and task difficulty on inattentional blindness. *Psychonomic Bulletin & Review*, *16*(2), 398–403. doi:<https://doi.org/10.3758/PBR.16.2.398>
- Tehrani, A. S. S., Lee, H., Mathews, S. C., Shore, A., Makary, M. A., Pronovost, P. J., & Newman-Toker, D. E. (2013). 25-Year summary of US malpractice claims for diagnostic errors 1986–2010: An analysis from the National Practitioner Data Bank. *BMJ Quality & Safety*, *22*(8), 672–680. doi:<https://doi.org/10.1136/bmjqs-2012-001550>
- Thomassen, Ø., Storesund, A., Søfteland, E., & Brattebø, G. (2014). The effects of safety checklists in medicine: A systematic review. *Acta Anaesthesiologica Scandinavica*, *58*(1), 5–18.
- Tuddenham, W. J. (1962). Visual Search, Image Organization, and Reader Error in Roentgen Diagnosis: Studies of the Psychophysiology of Roentgen Image Perception Memorial Fund Lecture. *Radiology*, *78*(5), 694–704.
- Ward, E. J., & Scholl, B. J. (2015). Inattentional blindness reflects limitations on perception, not memory: Evidence from repeated failures of awareness. *Psychonomic Bulletin & Review*, *22*(3), 722–727. doi:<https://doi.org/10.3758/s13423-014-0745-8>
- White, R. C., & Davies, A. A. (2008). Attention set for number: Expectation and perceptual load in inattentional blindness. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(5), 1092.
- Wolfe, J. M., Soce, A. A., & Schill, H. M. (2017). How did I miss that? Developing mixed hybrid visual search as a ‘model system’ for incidental finding errors in radiology. *Cognitive Research: Principles and Implications*, *2*.
- Wood, K., & Simons, D. J. (2019). The spatial allocation of attention in an interactive environment. *Cognitive Research: Principles and Implications*, *4*(1), 13. <https://doi.org/10.1186/s41235-019-0164-5>

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.