Trends in
Cognitive Sciences

CellPress

Review

Normal blindness: when we Look But Fail To See

Jeremy M. Wolfe,[1,2,*] Anna Kosovicheva,[3] and Benjamin Wolfe[3]

Humans routinely miss important information that is 'right in front of our eyes', from overlooking typos in a paper to failing to see a cyclist in an intersection. Recent studies on these 'Looked But Failed To See' (LBFTS) errors point to a common mechanism underlying these failures, whether the missed item was an unexpected gorilla, the clearly defined target of a visual search, or that simple typo. We argue that normal blindness is the by-product of the limited-capacity prediction engine that is our visual system. The processes that evolved to allow us to move through the world with ease are virtually guaranteed to cause us to miss some significant stimuli, especially in important tasks like driving and medical image perception.

Highlights

Looked But Failed to See (LBFTS) errors occur when observers fail to notice a clearly visible item, and occur across a wide range of tasks and settings, from driving and medical image perception to laboratory visual search tasks.

We outline a new, unified account of LBFTS errors, arguing that processes that serve us well under most circumstances are guaranteed to produce a steady stream of LBFTS errors under some circumstances.

LBFTS can be thought of as a form of 'normal blindness'. It is obviously far less severe than clinical blindness but it is so universal that its costs are substantial at a societal level.

When we miss what is right in front of our eyes

Most readers of *Trends in Cognitive Sciences* will be familiar with Simons and Chabris' classic gorilla experiments [1]. Observers watch two groups of actors passing a ball around. The observers are told to count the number of times that the white-shirted team passes the ball. In the middle of this game, another actor in a gorilla suit walks into the middle of the game, pounds on her chest, and leaves. When asked, after the video clip is done, observers can offer reasonable answers about the number of passes. However, even when asked in several different ways, about 50% of observers fail to report the gorilla. Those who missed it are surprised when shown the gorilla. How could they have been so 'blind'? It was clearly visible but, somehow, they did not see it. Now consider a different task. You dutifully proofread your latest paper, but it comes back later with several obvious typos marked. The misspellings were clearly visible, but you missed them. Finally, imagine doing a standard **visual search** (see Glossary) task; perhaps searching for a T among some number of Ls [2]. The stimuli are visible until the observer responds. Nevertheless, observers will typically reply that no target is present on 5–10% of trials containing a clearly visible 'T'. These all seem like different errors – what do they have in common? The gorilla task is a canonical example of **inattentional blindness** [3–5]. Problems with proofreading might be considered as 'satisfaction of search' errors [6–9] or, perhaps, vigilance failures [10–12]. Miss errors in visual search are often thought of in the context of the problem of search termination (i.e., when to end an unsuccessful search [13,14]). Although generally thought of and studied as different phenomena, each of these is an example of an LBFTS error [15]: a broad class of errors, defined by the failure to respond successfully to stimuli that are unambiguously visible and 'right in front of your eyes'.

LBFTS errors might occur in an 'incidental' setting [16], when observers are looking for or attending to something else, as in the gorilla example. Alternatively, these errors can occur in an 'intentional' setting, as when observers miss a perfectly visible T during an intentional search for a T. Some tasks lie in between. For instance, radiologists are asked to report 'incidental findings' that might be clinically significant [17]. For example, a radiologist, looking for lung cancer, commits a potentially dangerous LBFTS error if they miss visible evidence of pneumonia or vice versa [18,19]. In

[1]Brigham and Women's Hospital, 900 Commonwealth Avenue, Boston, MA 02215, USA
[2]Harvard Medical School, 25 Shattuck Street, Boston, MA 02115, USA
[3]Department of Psychology, University of Toronto Mississauga, 3359 Mississauga Road, Mississauga, Ontario, L5L 1C6, Canada

*Correspondence: jwolfe@bwh.harvard.edu (J.M. Wolfe).

 https://doi.org/10.1016/j.tics.2022.06.006

much more artificial situations, radiologists and others have been shown to miss gorillas in the context of performing their expert searches [20–22]. In the case of incidental findings, radiologists are specifically looking for signs of one pathology but know that there is a broader list of potential targets that might be present. By contrast, when a radiologist misses a gorilla in the lung [20], the situation is different because gorillas are not on the broader list.

Although these and other LBFTS situations seem like distinct phenomena, we argue, based on recent work, that they can all be seen as products of the same normal mechanisms of attention and object recognition. Specifically: (i) observers only select a subset of what they could process on each **fixation** (though they are not blind to the rest of the visual input); (ii) even the items that are selected by attention will be missed if too little time is given to their processing; (iii) the processes that give rise to routine visual awareness routinely persuade us that we have seen more than we have actually seen; and (iv) **attentional guidance** (attentional set) can guide observers away from targets as well as toward them. Taken together, these factors produce a state of 'normal blindness' that has significant implications for the way people function in the world.

Why do we care?

While missing a gorilla in an odd video or a typo in a manuscript are not serious errors, the family of LBFTS errors is broad and consequential. The term LBFTS comes from research on road collisions [15] (Box 1). In 'On-the-spot' interviews, after a collision drivers often insist that, yes, they did look at the location where the pedestrian or motorcyclist was located but, somehow, they just did not 'see' the victim (e.g., [23–25]). In medicine, radiologists routinely end up in court [26], sued for malpractice when they fail to report some clinically significant finding that is 'retrospectively visible' [27], meaning that the problem is visible when attention is drawn to it later. When people talk about 'blindness' in a clinical sense, they are usually referring to a devastating loss of visual function that has profound effects on the individual. By contrast, the 'normal blindness' that we are discussing here is not devastating at the individual level, because, by definition, it is part of 'normal' experience. However, because it is normal, the societal cost adds up. Developing a unified account of LBFTS errors and determining how to mitigate these errors promises to improve clinical outcomes, road safety, and, perhaps, even our ability to find typos.

Box 1. LBFTS errors in driving

The term 'looked but failed to see' is borrowed from road safety research, where it was coined to describe drivers' post-collision self-reports in which they reported that they had just not seen what they collided with [85]. These errors are not due to reduced visual acuity, impaired visual functioning, or reduced visual field extent [86]. Research on driver behavior was describing these failures as attentional in nature, long before the term LBFTS was coined [87], suggesting that purely visual accounts are insufficient to explain why drivers make these errors.

Research on LBFTS errors on the road emphasizes the role of expectation in causing these errors; a driver who does not expect to have to share the road with a motorcycle [88] or a cyclist [89] is less likely to see such fellow road users. Conversely, in environments where motorcycles or cyclists are common, drivers may be more likely to see them [90,91]. Interestingly, the driver's attentional set and level of expertise in driving do not seem to protect against LBFTS errors [24]. Drivers often multitask and distract themselves with other tasks while driving (from adjusting the radio to texting on a smartphone) [92] and the attentional and cognitive demands of these secondary tasks could magnify the risk of LBFTS errors even if hands are on the wheel and eyes are on the road [93]. These errors are closely linked with where drivers look and for how long [94], though applied research has focused on what drivers look at – to the exclusion of what is in the functional visual field (FVF), what is beyond it [35], or how drivers plan their eye movements [95]. Recently, the classic LBFTS account has been questioned, with researchers asking whether this phenomenon can be attributed to memory failures (i.e., a 'saw but forgot' error [96]).

Of course, LBFTS errors are not exclusive to driving. If researchers in multiple fields from medical image perception to basic cognitive science come to realize that LBFTS errors in one field are related to those in others, then the insights from different lines of research may converge on new ways to mitigate normal blindness.

Glossary

Asynchronous diffuser: information can accumulate in multiple 'diffusers' (see below) at the same time. If these processes start and stop at different times, the set of diffusers can be called an asynchronous diffuser.

Attentional guidance: the use of information from across the visual field to determine where the observer should attend next. For example, attention can be restricted to a subset of visual input based on basic features, location, or higher-level information (e.g., knowledge of scene structure). Also referred to as 'guidance'.

Diffuser: a process that accumulates information and outputs a response when the accumulating information reaches a threshold level.

Fixation: the location in the visual field that forms its image on the fovea (i.e., where the eyes are 'pointed').

Functional visual field (FVF): also called the Useful Field of View, the area around the point of fixation within which items can be (but might not be) recognized.

Hybrid search: a type of visual search task in which the observer needs to look for multiple different types of items (e.g., a radiologist looking for a cancerous nodule or pneumonia) within a single image.

Non-selective processes: visual processes that operate across the entire visual field allowing us to see something everywhere, even if selection and attention are required to actually identify what we are seeing.

Inattentional blindness: a failure to notice a clearly visible, but typically unexpected item in the absence of any visual impairment. Inattentional blindness can be understood as a Looked But Failed To See error.

Visual search: the process of looking for an item or multiple items (targets) among other objects (distractors).

A unifying model of LBFTS errors

Figure 1 offers a framework in which multiple types of LBFTS errors arise from the same underlying processes. To illustrate this model of LBFTS errors, it is useful to use a relatively complex task, as cartooned in Figure 1, as an example. Observers could be asked to find white 3s, a highly specific target, as well as any letters from the second half of the alphabet, a more categorical class of targets. The stimuli are presented in a noisy scene that, in this case, happens to be a mammogram. This is a **hybrid search** task, where participants are looking for more than one type of target [28]. Such a task is intended to mimic important properties of real-world tasks. For example, a driver might be attending to navigational cues, like signs and signals on the road and, at the same time, pointing out landmarks to the children. A radiologist might scan chest X-rays for signs of pneumonia while, at the same time, keeping watch for other 'incidental findings' that might be clinically significant, like cancerous nodules [29]. Drivers and radiologists are expected to be able to do this, and people blame them when they fail, thus it would be useful to have a model of why people make these errors that could suggest ways to prevent them.

We cannot fully process everything

Based on prior work, a story can be told about how this task might proceed. The eyes will be pointed at some location, for example, on the red plus in Figure 1A. When the eyes are directed at one location, there is a region surrounding that fixation within which a target can be identified. This is called the useful field of view (UFOV; [30]) or **functional visual field** (**FVF**; [31]) – the terms are roughly equivalent. UFOV is more common in the applied literature; FVF, in the basic, cognitive literature. The size of the FVF will vary with the target. Acuity limits and crowding [32] will mean that, to identify an object, the point of fixation needs to be much closer to a small target in a noisy scene than to a large item in an empty field [33].

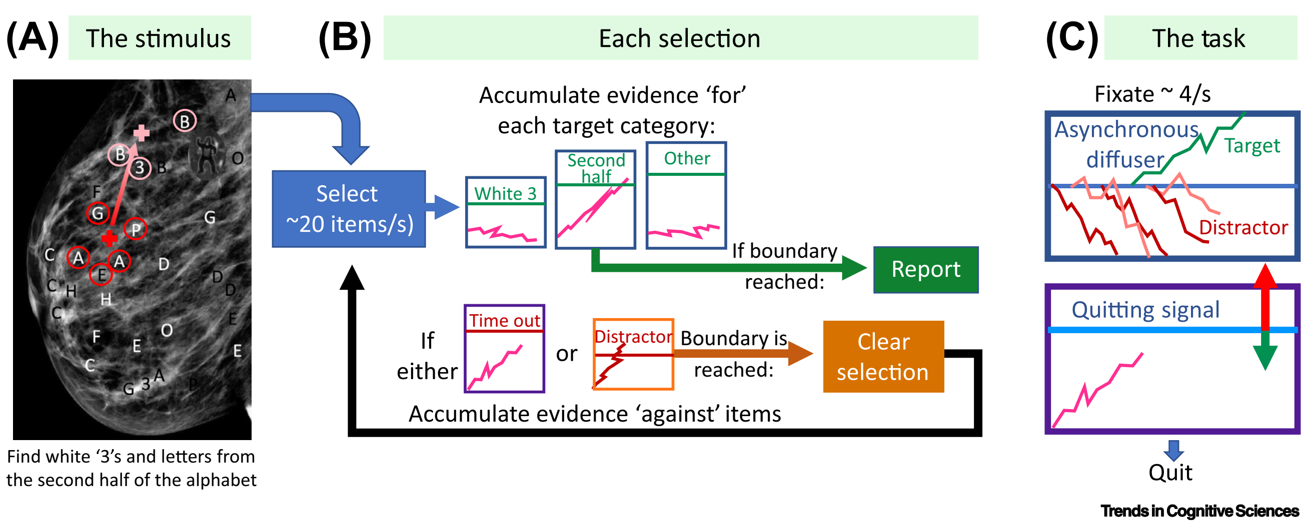

Figure 1. A model of normal visual processing that will produce a range of look but fail to see (LBFTS) errors. (A) Stimulus for a 'mixed hybrid search', where observers search for both specific targets (here, white 3s) and categorical targets (any letter from the second half of the alphabet). Plus signs show hypothetical fixation points. Circles show hypothetical deployments of attention. (B) Making a decision about a selected item may involve many information-accumulating processes (represented by each panel in the top row). One process seeks evidence for 'white 3s'. Another process, for letters in the second half of the alphabet, will take longer (higher decision boundary). Evidence for other 'surprises' would require still more evidence. Once enough evidence accumulates in one accumulator and the decision boundary is reached, the observer can report the target. If an item was identified as a distractor or if an item-by-item timing criterion was reached (bottom row), processing of that item would be terminated. At that point, the selection is cleared and a new item is selected for further processing. (C) At the level of the entire task, a limited number of items are selected into an asynchronous diffusion process, represented by the individual lines in the top panel. The task ends if the targets are found or when a task-level timer reaches a quitting threshold (bottom panel in C).

A standard, if often unstated, assumption is that people fully process whatever is inside the current FVF (e.g., [34]). This is incorrect [35]. To see why, consider the simpler search for Ts among Ls in Figure 2A. If you fixate on the red star, you will be able to identify the 'T' to the left. It is, by definition, inside the current FVF. Without moving your eyes, you will not be able to identify the T to the lower right. It is outside your FVF. If the left-hand T marks the outer bound of the current FVF, there would be something like 17 items within the FVF. A standard fixation lasts about 200–250 ms. Many lines of evidence tell us that you simply cannot fully process all of those 17 items in a quarter of a second. Perhaps the most straightforward evidence comes from standard visual search experiments in which observers search for targets like the T among Ls in Figure 2A. In a typical experiment, the set size (the number of items) varies from trial to trial. If we measure the response time (RT) to say 'yes, there is a target present' or 'no, there is not', we can calculate RT × set size functions. Those slopes estimate a rate of processing. The proper way to estimate the rate is the subject of debate [36–38]. Nevertheless, 40–50 items per second is about the highest rate that anyone would propose [39]. A rate of 20 items per second is more plausible for a T among L search. In a 200–250 ms fixation, that would suggest that 'at most' 8–12 of the 17 items in the FVF shown in Figure 2A could be processed to the point of identification. For present purposes, it does not matter how one envisions the details of processing within the FVF. It could involve serial selection of one item after another [40] or parallel but imperfect processing of all items [41] or some combination of serial and parallel processes [38,42]. The critical point is that it is not possible to fully process and identify everything within the FVF on a single fixation.

After ~200–250 ms, the eyes will move to a new location (e.g., the pink plus in Figure 1A) and the selection process will be repeated. The task, as a whole, has a structure as cartooned in Figure 1C. The search engine can be seen as an **asynchronous diffuser** [40,42] with information accumulating about several selected items at any one time. In addition, a task-level quitting mechanism [14,40,43] terminates searches that are not ended by the discovery of a target.

Inadequate processing of selected items

In Figure 1A, during that first fixation, attention might select the five items circled in red. When an item is selected, the system must decide what it is and if it is a target. Such decisions are

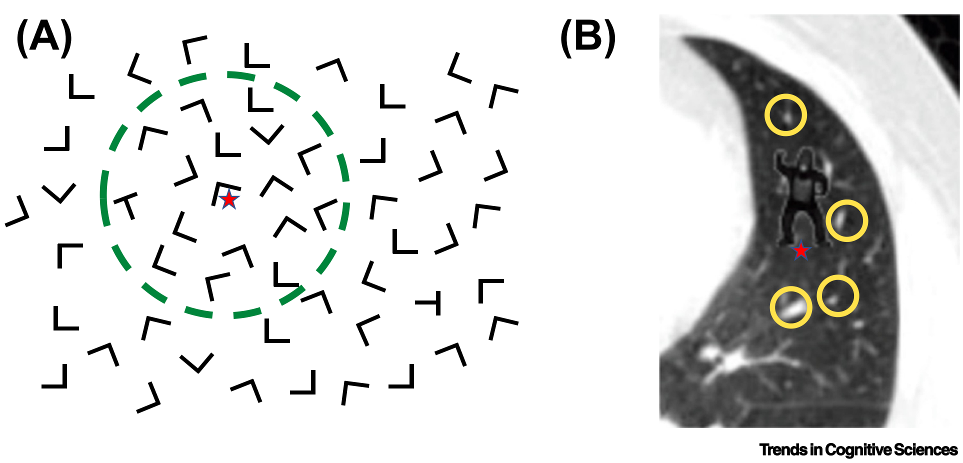

Figure 2. The role of the functional visual field (FVF) in look but fail to see (LBFTS) errors. (A) If observers fixated on the star, they would be able to identify the 'T' to the left but, in a quarter of a second fixation, they could only sample a subset of the items in the FVF (green circle) and might fail to select that T even if it is visible. (B) A section of a lung computed tomography (CT) image with an inserted gorilla. A radiologist, looking for small, white, round 'lung nodules' might be guided to select small white items while fixated on the star and, thus, be actively guided away from selecting and identifying the gorilla ('misguidance'). The missed T and missed gorilla LBFTS errors could both arise from the inability to fully process the contents of the FVF.

frequently modeled as diffusion processes where information accumulates in a noisy manner toward some decision boundary [44]. In talking about search, it is not often appreciated that recognition in all but the simplest tasks must involve multiple **diffusers** operating in parallel. For instance, look at the object in the lower left of Figure 3. Presumably, you identify 'bread rolls' when enough information accumulates to support that conclusion. At the same time, some weak evidence may accumulate for the conclusion that these are rocks, while little or no evidence supports the conclusion that this is, for instance, a rabbit. Moving to the right in the figure, it seems intuitively clear that you would be faster to accumulate the information needed to determine that the red thing is a car than to determine that it is a luxury item. Figure 1B shows one diffuser collecting evidence that the item is a 'white 3' target, another, that it is a target from the second half of the alphabet, and a third, making a decision about unlikely but incidentally interesting findings. Like the earlier examples of object identification (e.g., 'car' vs. 'luxury goods'), these will have different thresholds (and/or rates of information accumulation). It would be easier/faster to accumulate evidence that the selection is a specific item like a '3'. It will be harder/slower to confirm a letter in the second half of the alphabet, and still harder to reach the threshold for an unspecified but interesting 'other' item. Since only a small number of items can be selected, it is important to quickly reject useless selections. Figure 1B shows two ways this might happen. If the item is identified as a distractor (e.g., an 'A'), it can be rejected and space cleared in a limited-capacity workspace for a new selection. Hawkins and Heathcote's recent timed racing diffusion model

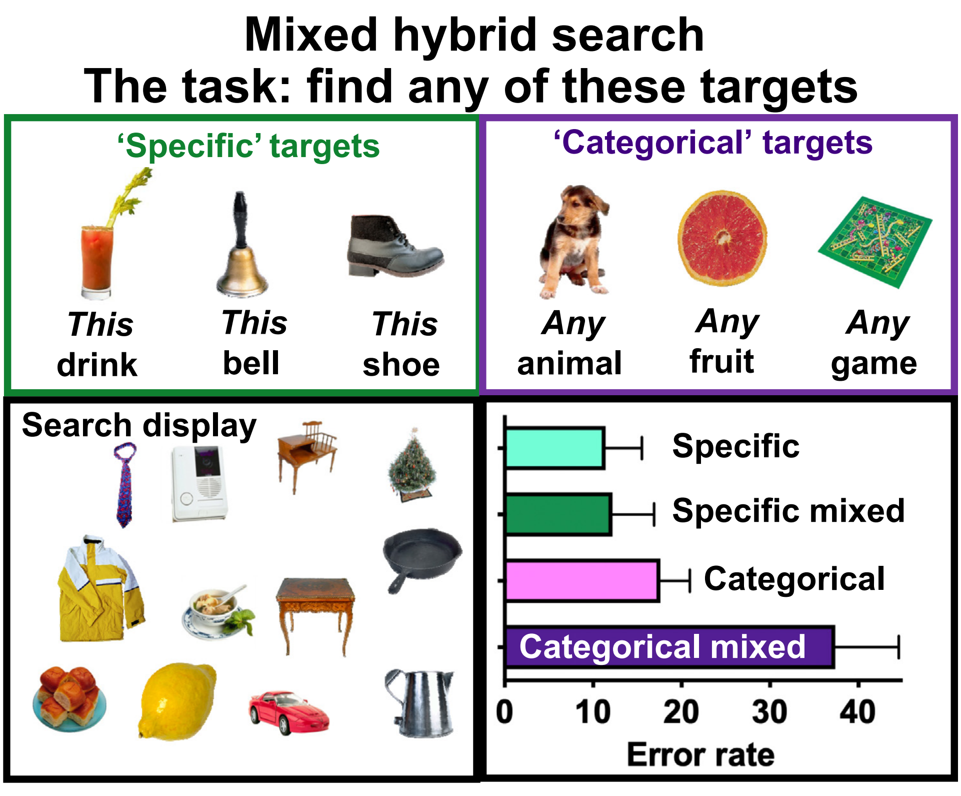

Figure 3. Mixed hybrid search. In the mixed hybrid search paradigm, observers search for a mix of specific targets (top-left: e.g., this specific drink) and categorical targets (top-right: e.g., any animal) within an array of distractor items (bottom-left). In these types of searches, observers frequently miss categorical targets (dark purple bar), especially if categorical targets are rare. The error rate is lower if all targets are categorical (light purple), and lower still for specific targets. The green bars show the error rates for specific targets when they appear together with categorical targets ('specific mixed') or on their own ('specific').

(TRDM) [45] argues for an important, second way to reject an item. Sometimes, no diffuser reaches a boundary in a reasonable amount of time. In that case, it must be possible to stop processing that item and move on to another one. The TRDM solution is to propose a time-based accumulator that ends processing for that item if its time boundary is reached (see also [46]). Imagine that you are trying to identify a bird. If you are not succeeding, it needs to be possible to disengage your attention and move on to other tasks. We would argue that it has not generally been recognized that, if only a few items can be selected for identification at one time, there must be a way to clear unprofitable selections and move on to other objects. Thus, there are two quitting mechanisms: one, to end processing a single item, and another to end the task as a whole.

Missing the designated target

The need to move on to the next item can explain some of the disturbing LBFTS errors where observers fail to see items that they were actually looking for. Consider the mixed hybrid search paradigm illustrated in Figure 3 [47]. In a hybrid search, observers search for any of several targets. In the mixed hybrid task depicted in Figure 3, targets were divided into two sets. There were three specific items (e.g., this drink, this bell, and this shoe) and three categorical targets (e.g., 'any' animal, 'any' fruit, 'any' game). The visual displays consisted of arrays of easily identifiable object photographs. As shown in the graph in Figure 3, the key finding is that observers missed many more of the categorical targets than the specific targets, especially on mixed blocks where they looked for a combination of specific and categorical targets. Importantly, in terms of the model in Figure 1, there is no reason to assume that categorical items would be fixated and/or selected less often than specific targets. Random selection would not favor the specific targets [48]. Something else must be contributing to the failure to identify a clearly visible lemon as a categorical fruit target. A hint is found in the RTs. It takes longer to successfully respond to categorical targets than to specific targets. It takes longer to decide that any item is an example of the category 'fruit' than to recognize that it is the correct, specific drink [49]. In Figure 1B terms, the categorical diffusers will have higher bounds and/or slower information accumulation.

The threshold for the item-by-item timer that Figure 1 borrowed from the TRDM [45] must be set for each task, presumably by some adaptive process that quits sooner if it can and quits later if too many mistakes are being made [13]. If the threshold is set too low, observers will flush items before identifying some more slowly identified, categorical items, even though they had been selected. This is a form of a speed–accuracy tradeoff [50,51]. If all the targets are categorical (light purple bar in the Figure 3 data graph), the item-quitting threshold will be set at a level producing a moderate number of errors. If specific items are mixed in, their faster identification will shift the quitting threshold to a shorter value, causing proportionally more categorical items to be missed. If specific targets are made markedly more common than categorical, observers will learn to move on even more quickly. Under such conditions, the data show that observers missed slower categorical targets at seven times the rate of specific targets, even though there is no reason to think that they were selected less often [47]. Thus, under the right circumstances, the same timing mechanism that normally speeds information processing can produce large numbers of LBFTS errors 'for targets that observers are actively searching for'. This path to LBFTS errors may be an important contributor to failures to report 'incidental findings' in medical images or to react to visible pedestrians in the roadway.

The role of target prevalence in LBFTS errors

In a mixed hybrid search, LBFTS errors increased as the probability (or 'prevalence') of the categorical targets declined. Target prevalence has a substantial effect on visual search tasks [52]. False-negative (miss) errors increase at low prevalence [53] making prevalence a plausible source of LBFTS errors in socially important, low-prevalence tasks like airport [54] and cancer screening

[55]. These errors, too, can fall out of the model in Figure 1. In search experiments, low prevalence makes decision criteria more conservative (fewer target-present responses) [56]. In a diffusion model like Figure 1, this is accomplished by moving the starting point of the accumulator closer to the distractor decision boundary and further from the target boundary for rare targets. Absent responses also become faster. This can be modeled as a change in the task-level-quitting mechanism in Figure 1C [56]. Thus, at the airport or the radiology clinic, an expert will require more evidence for the important low-prevalence targets and be more willing to give up on a search that has not yielded a target; again, a recipe for LBFTS errors (Box 2).

As a more mundane example of these mechanisms at work, think about the process of proofreading. Typos are rare as a fraction of all words. Most words serve as distractors in typo search, and the reader wants to quickly finish proofreading. Under these circumstances, the low-prevalence pressures will come into play and clearly identifiable misspellings will be missed.

Perceiving the most likely world

What do observers consciously experience as this process unfolds? Even if we can only select and attend to a subset of the items in the visual field, **non-selective processes** will deliver the impression of rapidly seeing something throughout the visual field [57,58]. While people are demonstrably not aware of the accurate details of the scene in the absence of selection and attention [59], the non-selective processes deliver a statistical summary of the current contents of a scene. This summary is often enough to create a reasonable and useful hypothesis about what is present [60–63], even if it is not enough to tell us if a 3 (or a breast nodule) is present or where it is. In addition, observers do not have introspective access to what occurs at the timescale of attentional deployments, lumping perhaps several hundred milliseconds into a rough sense of 'now' [64,65]. People are also surprisingly bad at knowing where they have looked [66]. Indeed, patients with visual deficits can be surprisingly bad at knowing where in the visual field they are impaired. For example, patients with significant scotomata may assert that the visual field appears not to have any holes in it [67] and, of course, people do not see their normal physiological blindspots [68].

But what do we 'see'? We do not see a mass of image statistics. It seems that observers experience their best guess about the state of the world based on what Helmholtz called 'unconscious

Box 2. Prevalence effects: LBFTS errors for rare targets

Classic studies of vigilance, dating back to the 1940s, show that when observers monitor for rare, but important signals (e.g., a target on a radar screen), these events are frequently missed [97]. Although these targets are visible and easily detected under normal circumstances, observers miss them after a prolonged period of sustained attention. Much of this work has attributed these misses to a loss in sensitivity (i.e., a reduced ability to distinguish signal events from noise) [98]. However, more recent findings challenge this account, attributing the vigilance decrement to a shift in decision boundaries, or criteria [99]. In other words, as observers monitor for rare events over a long period, they may be less willing to say that a target is present.

This result parallels those seen in visual search tasks. When observers look for rare items, like an abnormality in a CT scan, or threats in luggage [100], these targets are often missed. As in the vigilance literature, this increase in misses can be attributed to a shift in decision boundaries or criteria [101]. In other words, our perceptual ability to detect the target is the same, but as they become rare, we become less willing to say that something is there. Many efforts have been made to 'cure' this prevalence effect, with mixed success [102–105]. To what extent are classic vigilance failures and low-prevalence search errors the same type of LBFTS error? Both seem to involve changes in decision criteria, though typical vigilance errors involve transient stimuli, while LBFTS errors in search usually involve stimuli that are present until the observer dismisses them. Moreover, while vigilance tasks generally involve monitoring for a single event, prevalence effects in visual search involve a series of decisions (i.e., is this item a T or an L? What about the next one?), along with an overall decision about when to stop searching (i.e., a task-level quitting threshold). Models have been proposed for prevalence effects in search accounting for both types of decisions [14,56,106] and vigilance [11]. An integrated model would be useful.

inference' [69]. A more modern, Bayesian 'predictive coding' account would say that 'the hypothesis with the highest posterior probability (i.e., most probable given the input) wins and gets to determine the perceptual content of the system.' [70]; (for a review see [71]). For present purposes, the important point is that after fixating on a location for a quarter of a second, observers are likely to be convinced that they saw and understood whatever was there to be seen, even if their actions and the data can clearly show that this was not the case. Moreover, an important aspect of maintaining a plausible theory about the state of the world is what Friston [72] describes as 'suppressing surprise'. In terms of the model in Figure 1, this would correspond to raising decision bounds on unlikely accumulators so that the system does not jump to unlikely conclusions about a selected item. For example, observers should not conclude that an 'A' is the Matterhorn peak, even though some signal to that effect might accumulate. Of course, the system should be tuned so that it can change its current hypothesis about the world in the face of enough surprising data, but a high decision boundary for unlikely items decreases the probability that a radiologist will report an incidental finding or that a driver will notice a pedestrian in an unlikely location.

We have described the role of three of the four factors, listed at the outset: (i) during a fixation, a subset of all possible objects of attention will be selected for limited-capacity processing; (ii) if a selected item does not yield its identity reasonably quickly, it will be flushed to make space for a new, perhaps more promising, source of information [73]; and (iii) the observer will experience the currently most probable state of the world, avoiding most false alarms. That observer will be surprised when they miss a clearly visible item because they will not realize that they never selected it or that they selected it but moved on before fully analyzing it. One more factor helps us to understand the phenomena like the missed gorilla in the opening example.

Missing an unexpected target: guidance and misguidance

In the case of a T among L search as in Figure 2A, observers probably process a fairly random subset of items, with, perhaps, a bias toward selecting items closer to fixation [74]. In other tasks, including the task in Figure 1, the choice of what to process is unlikely to be random. In the Figure 1A task, attention would be guided to white items in the search for a white 3 or, in a version of the Figure 2A task, imagine observers are looking for red Ts among black and red Ls. They will not attend to many black Ls. Attention will be guided by basic feature information [75], in this case, to the red items [76].

Normally, guidance is very useful. In some instances, like the classic gorilla study, guidance becomes misguidance. In that study, observers monitored the white-shirted team. In another version, they were less likely to miss the gorilla if they monitored the black-shirted team (see also [77]). If they were guided to black shirts, they were more likely to notice a black gorilla. Figure 2B illustrates this point using a piece of a lung computed tomography (CT) image from an inattentional blindness experiment with radiologists as the participants [20]. The radiologists were searching for small, white, round lung nodules and missed a large, black, irregular but easily detectable gorilla inserted into the image. Eye-tracking shows that the radiologists almost always fixated on or near the gorilla. As shown in Figure 2B, a radiologist hypothetically fixated on the red star might be guided, as a result of their expertise, toward four to five plausible lung nodules and thus be guided away from the gorilla. Even if the gorilla was selected, its processing might have been terminated before the gorilla was identified because, in keeping with the efforts of a predictive coding process to avoid false alarm surprises, the threshold for identifying an item as unlikely as a gorilla would be set high. After all, gorillas just do not happen in the lung or, for that matter, in the midst of ball passing games [1]. Thus, 'misguidance' and the need to end the processing of selections that do not produce rapid identifications can give rise to inattentional blindness. By the

Box 3. Can the computer save us from LBFTS errors?

'Curing' normal blindness by focusing on the observer may be hard [101]. Training radiologists or pilots to use a certain gaze pattern to scan a mammogram [20] or the view out the windscreen [107] may help, but is only part of the solution [108]. Perhaps computer algorithms that do not forget to do what they are told could be the answer to LBFTS errors, because they can tell us what we might miss. Your car can tell you when a child runs into the road [109], or if you are a radiologist, it can tell you that this region of a mammogram looks suspicious [110]. Large investments have led to significant advances in computer-aided detection and more progress will be coming. That said, the ubiquitous spellchecker alerts us to possible problems:

(i) While it can be amusing when the spellchecker offers up the wrong word, if the 'positive predictive value' (proportion of correct suggestions out of all suggestions) is not high, users will tend to ignore an AI system [111]. Similarly, a driver may be less willing to trust an autonomous driving system if it is inconsistent or unreliable (see [112] for a review).
(ii) At the other extreme, if the system is very reliable, the user may lose the skill required to recover in those cases where the AI fails. For example, how are your map reading skills now that you are regularly taking direction from your phone? [113]. This kind of overreliance or 'de-skilling' has long been a concern for pilots and is often a question in aviation accidents [114].

way, if you failed to note the gorilla in Figure 1A, you can search for it now. It should be visible. Misguidance is central to many effective magic tricks as well [78,79].

Concluding remarks

Thus, from gorillas in the lung, to a cyclist on the road, to the lowly typo, a constellation of LBFTS errors can be accounted for as the by-product of quite reasonable settings of the parameters of a mechanism like that shown in Figure 1. Some of the LBFTS components, reviewed here, have been suggested previously. Here, we unite them into a single framework. In previous discussions, there has been less emphasis on the idea that an item could be attended to and still missed. The mixed hybrid result (Figure 3) makes it clear that this must happen. The novel, item-by-item quitting mechanism in Hawkins and Heathcote's TRDM [45] offers a potential mechanism by which items can be missed, even if attended.

In our view, LBFTS errors are not the result of doing something wrong. They are the product of situations where the normal functioning of the visual system produces surprising or undesirable outcomes. We refer to this as 'normal blindness' to emphasize the idea that LBFTS errors are a form of visual impairment. Unlike clinical blindness, everyone experiences normal blindness and people are mostly blind to its costs (normal blindness blindness? [80]). Normal blindness is probabilistic. An item that is missed by one observer at one time will be seen at a different time or by a different observer. This guarantees, for instance, that a continuity error, missed by a filmmaker, will be noted and reported by some of the many observers who later see the film [16]. What can be done? (see Outstanding questions). Given that errors can arise from statistical bad luck, a second pair of eyes can help (i.e., who found those typos in your paper, after all.) That second reader might be a computer (e.g., [81,82], (Box 3). In settings like driving, we can try to build an environment that helps us use our limited resources [83,84]. Since people do not want to continue paying the costs of normal blindness in lives lost in medicine or on the road (or even in typos missed), having a model of normal blindness gives us some of the tools we need to understand it and work to reduce it. If we all suffer from normal blindness, we should work to minimize LBFTS errors because the result will benefit all of us.

Outstanding questions

Clearly, the central question is, 'How can we reduce Normal Blindness?' As we argue here, it may be based in mechanisms of visual attention and not easily resolved through training or awareness (e.g., by informing observers of these errors). Can LBFTS errors be reduced through manipulations that implicitly or explicitly shift the observer's decision boundary, perhaps through changing the perceived costs and payoffs?

To what extent can technology help? (Artificial intelligence in medicine, self-driving cars, etc.)

Do motivated observers make fewer errors? Surely, the answer is 'yes' but how far can motivation go in eliminating errors? More broadly, how do LBFTS errors vary with the mental state of the observer (e.g., are these errors more likely under fatigue)?

Similarly, are some individuals more susceptible to these errors? For example, how do LBFTS errors compare between younger and older adults? Are there particularly dangerous interactions of normal blindness with clinical visual or cognitive impairments?

Observers often experience LBFTS errors as 'blindness' – I did not *see* it. How much information about the unseen target is, in fact, processed outside of awareness, and to what extent can any of the subconsciously processed information affect observers' actions? Would you walk around a gorilla that you could not report seeing?

The model discussed here focuses on visual/attentional factors. How much of the LBFTS problem is a memory problem? The observer fails to recall the whole list of possible targets or, alternatively, they remember all the targets but do not successfully load them into working memory or activated long-term memory.

Acknowledgments

This work was funded by National Institutes of Health (NIH) grants CA207490 and EY017001 to J.M.W. The authors thank Farah Wick, Daniel Ernst, Wanyi Liu, and Jan Theeuwes for comments.

Declaration of interests

No interests are declared.

References

1. Simons, D.J. and Chabris, C.F. (1999) Gorillas in our midst: sustained inattentional blindness for dynamic events. *Perception* 28, 1059–1074
2. Wolfe, J.M. (2020) Visual search: how do we find what we are looking for? *Annu. Rev. Vis. Sci.* 6, 539–562
3. Mack, A. and Rock, I., eds (1998) *Inattentional Blindness*, MIT Press
4. Simons, D.J. (2000) Attentional capture and inattentional blindness. *Trends Cogn. Sci.* 4, 147–155
5. Rensink, R.A. (2000) Seeing, sensing, and scrutinizing. *Vis. Res.* 40, 1469–1487
6. Tuddenham, W.J. (1962) Visual search, image organization, and reader error in roentgen diagnosis. Studies of the psycho-physiology of roentgen image perception. *Radiology* 78, 694–704
7. Nodine, C.F. *et al.* (1992) Satisfaction of search (SOS). *Investig. Radiol.* 27, 571–573
8. Berbaum, K.S. *et al.* (2010) Satisfaction of search in traditional radiographic imaging. In *The Handbook of Medical Image Perception and Techniques* (Krupinski, E.A. and Samei, E., eds), pp. 107–138, Cambridge University Press
9. Barach, E. *et al.* (2021) Satisfaction-of-search (SOS) impacts multiple-target searches during proofreading: evidence from eye movements. *Vis. Cogn.* 29, 510–518
10. Warm, J. *et al.* (2014) Vigilance: a perceptual challenge. In *The Cambridge Handbook of Applied Perception Research* (Hoffman, Robert R. *et al.*, eds), Cambridge University Press
11. Thomson, D.R. *et al.* (2015) A resource-control account of sustained attention: evidence from mind-wandering and vigilance paradigms. *Perspect. Psychol. Sci.* 10, 82–96
12. McCarley, J.S. and Yamani, Y. (2021) Psychometric curves reveal three mechanisms of vigilance decrement. *Psychol. Sci.* 32, 1675–1683
13. Chun, M.M. and Wolfe, J.M. (1996) Just say no: How are visual searches terminated when there is no target present? *Cogn. Psychol.* 30, 39–78
14. Schwarz, W. and Miller, J.O. (2016) GSDT: an integrative model of visual search. *J. Exp. Psychol. Hum. Percept. Perform.* 42, 1654–1675
15. Hills, B.L. (1980) Vision, visibility, and perception in driving. *Perception* 9, 183–216
16. Simons, D.J. and Mitroff, S.R. (2001) The role of expectations in change detection and attentional capture. In *Vision and Attention* (Jenkin, M. and Harris, L., eds), pp. 189–207, Springer
17. Moore, C.L. *et al.* (2021) Incidental findings: a survey of radiologists and emergency physicians. *J. Am. Coll. Radiol.* 18, 853–856
18. Beigelman-Aubry, C. *et al.* (2007) Management of an incidentally discovered pulmonary nodule. *Eur. Radiol.* 17, 449–466
19. Wu, C.C. *et al.* (2012) Incidental pulmonary nodules detected on abdominal computed tomography. *J. Comput. Assist. Tomogr.* 36, 641–645
20. Drew, T. *et al.* (2013) The invisible gorilla strikes again: sustained inattentional blindness in expert observers. *Psychol. Sci.* 24, 1848–1853
21. Williams, L. *et al.* (2021) The invisible breast cancer: experience does not protect against inattentional blindness to clinically relevant findings in radiology. *Psychon. Bull. Rev.* 28, 503–511
22. Gascho, D. *et al.* (2014) The invisible gorilla! A visual note of inattentional blindness. *J. Forensic Radiol. Imaging.* 2, 221
23. Clabaux, N. *et al.* (2012) Motorcyclists' speed and "looked-but-failed-to-see" accidents. *Accid. Anal. Prev.* 49, 73–77
24. Herslund, M.-B. and Jørgensen, N.O. (2003) Looked-but-failed-to-see-errors in traffic. *Accid. Anal. Prev.* 35, 885–891
25. Langham, M. *et al.* (2002) An analysis of 'looked but failed to see' accidents involving parked police vehicles. *Ergonomics* 45, 167–185
26. Berlin, L. (2007) Radiologic errors and malpractice: a blurry distinction. *Am. J. Roentgenol.* 189, 517–522
27. Nodine, C.F. *et al.* (2001) Blinded review of retrospectively visible unreported breast cancers: an eye-position analysis. *Radiology* 221, 122–129
28. Schneider, W. and Shiffrin, R.M. (1977) Controlled and automatic human information processing: I. Detection, search, and attention. *Psychol. Rev.* 84, 1–66
29. Schlett, C.L. *et al.* (2021) Incidental findings in whole-body MR imaging of a population-based cohort study: frequency, management and psychosocial consequences. *Eur. J. Radiol.* 134, 109451
30. Ball, K.K. *et al.* (1988) Age and visual search: expanding the useful field of view. *J. Opt. Soc. Am. A.* 5, 2210–2219
31. Sanders, A.F. (1970) Some aspects of the selective process in the functional visual field. *Ergonomics* 13, 101–117
32. Whitney, D. and Levi, D.M. (2011) Visual crowding: a fundamental limit on conscious perception and object recognition. *Trends Cogn. Sci.* 15, 160–168
33. Hulleman, J. and Olivers, C.N.L. (2017) The impending demise of the item in visual search. *Behav. Brain Sci.* 40, e132
34. Andersen, G.J. *et al.* (2011) Limits of spatial attention in three-dimensional space and dual-task driving performance. *Accid. Anal. Prev.* 43, 381–390
35. Wolfe, B. *et al.* (2017) More than the useful field: considering peripheral vision in driving. *Appl. Ergon.* 65, 316–325
36. Ward, R. *et al.* (1996) The slow time-course of visual attention. *Cogn. Psychol.* 30, 79–109
37. Wolfe, J.M. *et al.* (2000) Attention is fast but volition is slow. *Nature* 406, 691
38. Liesefeld, H. and Mueller, H.J. (2020) A theoretical attempt to revive the serial/parallel-search dichotomy. *Atten. Percept. Psychophys.* 82, 228–245
39. Wu, C.C. and Wolfe, J.M. (2022) The functional visual field(s) in simple visual search. *Vis. Res.* 190, 107965
40. Wolfe, J.M. (2021) Guided Search 6.0: an updated model of visual search. *Psychon. Bull. Rev.* 28, 1060–1092
41. Young, A.H. and Hulleman, J. (2013) Eye movements reveal how task difficulty moulds visual search. *J. Exp. Psychol. Hum Percept. Perform.* 39, 168–190
42. Wolfe, J.M. (2003) Moving towards solutions to some enduring controversies in visual search. *Trends Cogn. Sci.* 7, 70–76
43. Moran, R. *et al.* (2013) Competitive guided search: meeting the challenge of benchmark RT distributions. *J. Vis.* 13, 24
44. Ratcliff, R. (1978) A theory of memory retrieval. *Psych. Rev.* 85, 59–108
45. Hawkins, G.E. and Heathcote, A. (2021) Racing against the clock: evidence-based versus time-based decisions. *Psychol. Rev.* 128, 222–263
46. Ackerman, R. (2014) The diminishing criterion model for metacognitive regulation of time investment. *J. Exp. Psychol. Gen.* 143, 1349–1368
47. Wolfe, J.M. *et al.* (2017) How did I miss that? Developing mixed hybrid visual search as a 'model system' for incidental finding errors in radiology. *Cogn. Res. Princ. Implic. (CRPI)* 2, 35
48. Vickery, T.J. *et al.* (2005) Setting up the target template in visual search. *J. Vis.* 5, 81–92
49. Ibsen, C. *et al.* (1972) Fictive kin term use and social relationships – alternative interpretations. *J. Marriage Fam.* 34, 615–620
50. Henmon, V.A.C. (1911) The relation of the time of a judgment to its accuracy. *Psychol. Rev.* 18, 186–201
51. Heitz, R.P. (2014) The speed-accuracy tradeoff: history, physiology, methodology, and behavior. *Front. Neurosci.* 8, 150
52. Horowitz, T.S. (2017) Prevalence in visual search: from the clinic to the lab and back again. *Jpn. Psychol. Res.* 59, 65–108
53. Mitroff, S.R. and Biggs, A.T. (2014) The ultra-rare item effect: exceedingly rare items are highly susceptible to search errors. *Psychol. Sci.* 25, 284–289
54. Wolfe, J.M. *et al.* (2013) Prevalence effects in newly trained airport checkpoint screeners: Trained observers miss rare targets, too. *J. Vis.* 13, 33
55. Evans, K.K. *et al.* (2013) 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, e64366

56. Wolfe, J.M. and VanWert, M.J. (2010) Varying target prevalence reveals two, dissociable decision criteria in visual search. *Curr. Biol.* 20, 121–124
57. Potter, M.C. and Faulconer, B.A. (1975) Time to understand pictures and words. *Nature* 253, 437–438
58. Oliva, A. and Torralba, A. (2001) Modeling the shape of the scene: a holistic representation of the spatial envelope. *Int. J. Comp. Vis.* 42, 145–175
59. O'Regan, K. (1992) Solving the 'real' mysteries of visual perception. The world as an outside memory. *Can. J. Psychol.* 46, 461–488
60. Ariely, D. (2001) Seeing sets: representation by statistical properties. *Psychol. Sci.* 12, 157–162
61. Chong, S.C. and Treisman, A. (2003) Representation of statistical properties. *Vis. Res.* 43, 393–404
62. Alvarez, G.A. (2011) Representing multiple objects as an ensemble enhances visual cognition. *Trends Cogn. Sci.* 15, 122–131
63. Rosenholtz, R. (2020) Demystifying visual awareness: peripheral encoding plus limited decision complexity resolve the paradox of rich visual experience and curious perceptual failures. *Atten. Percept. Psychophys.* 82, 901–925
64. Wittmann, M. (2011) Moments in time. *Front. Integr. Neurosci.* 5, 66
65. Holcombe, A.O. (2009) Seeing slow and seeing fast: two limits on perception. *Trends Cogn. Sci.* 13, 216–221
66. Vo, M.L. *et al.* (2016) You think you know where you looked? You better look again. *J. Exp. Psychol Hum. Percept. Perform.* 42, 1477–1481
67. Schuchard, R.A. (1995) Adaptation to macular scotomas in persons with low vision. *Am. J. Occup. Ther.* 49, 870–876
68. Ramachandran, V.S. (1992) Blind spots. *Sci. Am.* 266, 86–91
69. von Helmholtz, H. (1924) *Treatise on Physiological Optics (Trans. from 3rd German ed. of 1909, edn.)*, The Optical Society of America
70. Hohwy, J. *et al.* (2008) Predictive coding explains binocular rivalry: an epistemological review. *Cognition* 108, 687–701
71. Kersten, D. *et al.* (2004) Object perception as Bayesian inference. *Annu. Rev. Psychol.* 55, 271–304
72. Friston, K. (2009) The free-energy principle: a rough guide to the brain? *Trends Cogn. Sci.* 13, 293–301
73. Najemnik, J. and Geisler, W.S. (2005) Optimal eye movement strategies in visual search. *Nature* 434, 387–391
74. Carrasco, M. *et al.* (1995) The eccentricity effect: target eccentricity affects performance on conjunction searches. *Percept. Psychophys.* 57, 1241–1261
75. Wolfe, J.M. and Horowitz, T.S. (2017) Five factors that guide attention in visual search. *Nat. Hum. Behav.* 1, 0058
76. Egeth, H.E. *et al.* (1984) Searching for conjunctively defined targets. *J. Exp. Psychol. Hum. Percept. Perform.* 10, 32–39
77. Most, S.B. *et al.* (2001) How not to be seen: the contribution of similarity and selective ignoring to sustained inattentional blindness. *Psychol. Sci.* 12, 9–17
78. Kuhn, G. *et al.* (2008) Towards a science of magic. *Trends Cogn. Sci.* 12, 349–354
79. Kuhn, G. *et al.* (2009) You look where I look! Effect of gaze cues on overt and covert attention in misdirection. *Vis. Cogn.* 17, 925–944
80. Levin, D.T. *et al.* (2000) Change blindness: the metacognitive error of overestimating change-detection ability. *Vis. Cogn.* 7, 397–412
81. Nam, J.G. *et al.* (2020) Undetected lung cancer at posteroanterior chest radiography: potential role of a deep learning–based detection algorithm. *Radiol. Cardiothoracic Imaging* 2, e190222
82. McKinney, S.M. *et al.* (2020) International evaluation of an AI system for breast cancer screening. *Nature* 577, 89–94
83. Theeuwes, J. (2021) Self-explaining roads: What does visual cognition tell us about designing safer roads? *Cogn. Res. Princ. Implic.* 6, 15
84. Theeuwes, J. and Godthelp, H. (1995) Self-explaining roads. *Saf. Sci.* 19, 217–225
85. Hills, T.T. (2006) Animal foraging and the evolution of goal-directed cognition. *Cogn. Sci.* 30, 3–41
86. Burg, A. (1971) Vision and driving: a report on research. *Hum. Factors* 13, 79–87
87. Gibson, J.J. and Crooks, L.E. (1938) A theoretical field-analysis of automobile-driving. *Am. J. Psychol.* 51, 453–471
88. Crundall, D. *et al.* (2012) Why do car drivers fail to give way to motorcycles at t-junctions? *Accid. Anal. Prev.* 44, 88–96
89. Kaya, N. *et al.* (2021) Hey, watch where you're going! An on-road study of driver scanning failures towards pedestrians and cyclists. *Accid. Anal. Prev.* 162, 106380
90. Beanland, V. *et al.* (2014) Safety in numbers: target prevalence affects the detection of vehicles during simulated driving. *Atten. Percept. Psychophys.* 76, 805–813
91. Jacobsen, P.L. (2015) Safety in numbers: more walkers and bicyclists, safer walking and bicycling. *Inj. Prev.* 21, 271–275
92. Strayer, D.L. and Cooper, J.M. (2015) Driven to Distraction. *Hum. Factors* 57, 1343–1347
93. White, C.B. and Caird, J.K. (2010) The blind date: the effects of change blindness, passenger conversation and gender on looked-but-failed-to-see (LBFTS) errors. *Accid. Anal. Prev.* 42, 1822–1830
94. Crundall, D. *et al.* (2012) Some hazards are more attractive than others: drivers of varying experience respond differently to different types of hazard. *Accid. Anal. Prev.* 45, 600–609
95. Wolfe, B. *et al.* (2020) Toward a theory of visual information acquisition in driving. *Hum. Factors* 64, 694–713
96. Robbins, C.J. *et al.* (2019) The 'saw but forgot' error: a role for short-term memory failures in understanding junction crashes? *PLoS One* 14, e0222905
97. Mackworth, N.H. (1948) The breakdown of vigilance during prolonged visual search. *Q. J. Exp. Psychol.* 1, 6–21
98. See, J.E. *et al.* (1995) Meta-analysis of the sensitivity decrement in vigilance. *Psychol. Bull.* 117, 230–249
99. Thomson, D.R. *et al.* (2016) A critical examination of the evidence for sensitivity loss in modern vigilance tasks. *Psychol. Rev.* 123, 70–83
100. Wolfe, J.M. *et al.* (2005) Rare targets are often missed in visual search. *Nature* 435, 439–440
101. Wolfe, J.M. *et al.* (2007) Low target prevalence is a stubborn source of errors in visual search tasks. *J. Exp. Psychol. Gen.* 136, 623–638
102. Kunar, M.A. *et al.* (2010) Spatial and temporal separation fails to counteract the effects of low prevalence in visual search. *Vis. Cogn.* 18, 881–897
103. Lau, J.S. and Huang, L. (2010) The prevalence effect is determined by past experience, not future prospects. *Vis. Res.* 50, 1469–1474
104. Reed, W.M. *et al.* (2011) The effect of abnormality-prevalence expectation on expert observer performance and visual search. *Radiology* 258, 938–943
105. Taylor, E. *et al.* (2022) Eliminating the low prevalence effect in visual search with a remarkably simple strategy. *Psych. Sci.* 33, 716–724
106. Trueblood, J.S. *et al.* (2021) Disentangling prevalence induced biases in medical image decision-making. *Cognition* 212, 104713
107. Ahmadi, N. *et al.* (2022) Improving the tactical scanning of student pilots: a gaze-based training intervention for transition from visual flight into instrument meteorological conditions. *Appl. Ergon.* 100, 103642
108. Ashraf, H. *et al.* (2018) Eye-tracking technology in medical education: a systematic review. *Med. Teach.* 40, 62–69
109. Källhammer, J.-E. *et al.* (2017) Modeling ratings of in-vehicle alerts to pedestrian by leveraging field operational tests data in a controlled laboratory study. *Transport. Res. F: Traffic Psychol. Behav.* 46, 413–425
110. Brem, R.F. *et al.* (2003) Improvement in sensitivity of screening mammography with computer-aided detection: a multiinstitutional trial. *Am. J. Roentgenol.* 181, 687–693
111. Nishikawa, R.M. and Bae, K.T. (2018) Importance of better human-computer interaction in the era of deep learning: mammography computer-aided diagnosis as a use case. *J. Am. Coll. Radiol.* 15, 49–52
112. Manchon, J.B. *et al.* (2021) From manual to automated driving: how does trust evolve? *Theor. Issues Ergon. Sci.* 22, 528–554
113. Bachmann, M.M. (2015) Exploring and addressing functional map illiteracy in community college students. *Pap. Appl. Geogr.* 1, 258–261
114. Bainbridge, L. (1983) Ironies of automation. *Automatica* 19, 775–779