The flash-preview moving window paradigm: Unpacking visual expertise one glimpse at a time

Damien Litchfield[1] & Tim Donovan[2]

[1] Edge Hill University, UK

[2] University of Cumbria, UK

Abstract

How we make sense of what we see and where best to look is shaped by our experience, our current task goals and how we first perceive our environment. An established way of demonstrating these factors work together is to study how eye movement patterns change as a function of expertise and to observe how experts can solve complex tasks after only very brief glances at a domain-specific image. The primary focus of this paper is to introduce an innovative gaze-contingent method called the 'Flash-Preview Moving Window' (FPMW) paradigm (Castelhano & Henderson, 2007), which was recently developed to understand our shared expertise in scene perception and how our first glimpse of a scene is used to guide our eye movement behaviour. In keeping with this special issue on visual expertise and medicine, this paper will highlight how the FPMW paradigm has the potential to resolve long-standing theoretical issues as to how, right from the very first glance, experts are able to process domain-specific images and guide their eye movements better than novices. Since FPMW is a gaze-contingent eye-tracking method, the paper will first outline the current methodological and theoretical frontier, and how the FPMW paradigm bridges established methods used to investigate visual expertise. The paper will discuss a recent example in which the FPMW was employed to investigate medical image perception expertise for the first time (Litchfield & Donovan, 2016), and by discussing the insights and challenges this method offers, this should ultimately deepen our understanding of visual expertise.

Keywords: flash-preview moving window; eye movements; medical image perception; visual expertise; eye-tracking

1. Introduction

From the moment we open our eyes we see a rich visual world. We are unable to process all of this incoming information and so we must move our eyes several times a second to look at and process different aspects of our environment. How we make sense of what we see and where best to look is shaped by our experience, our current task goals and how we first perceive our environment (Buswell, 1935; Henderson, 2007; Rayner, 2009; Yarbus, 1967). The common purpose of most eye tracking studies is to explore how eye movement behaviour is regulated and to see how eye movement behaviour during specific tasks relates to underlying visual and cognitive processes (Just & Carpenter, 1984). When searching for an object in a newly presented image, the first fixation quickly (i.e., within 40-100ms) encapsulates the initial "gist" of the scene and receives some pre-attentive processing of basic features (Rayner, Smith, Malcom, & Henderson, 2009; Võ & Henderson, 2010). At this point visual properties such as colour, contour distribution and spatial frequency are likely to be processed (Henderson & Hollingworth, 1999; Oliva, & Schyns, 1997; Schyns & Oliva, 1994), with scene context and semantic information accessible depending on the duration of the first glimpse of the scene (Fei-Fei, Iyer, Koch, & Perona, 2007; Potter, 1976). It is within this initial glimpse of the scene that subsequent eye movements are guided (Castelhano & Henderson, 2007) with parafoveal and peripheral vision playing an important role in the early comprehension of the gist of a scene and in the detection of targets (Henderson, Pollatsek, & Rayner, 1989). The visual system is able to integrate this visual input from sensory information with top-down information, which allows us to identify and locate any task-relevant item(s) within the receptive field for saccadic targeting. However, the processes that underlie this integration of information are still heavily debated (see Cohen, Dennett, & Kanwisher, 2016; Tatler, 2009; Torralba, Oliva, Castelhano, & Henderson, 2006; Wolfe Evans, Võ, & Greene, 2011; Zelinsky & Schmidt, 2009). Understanding how these visual processes become optimised with experience not only sheds light on these cutting edge issues, but also provides new ways of studying the development (and potential enhancement) of expertise (Donovan & Litchfield, 2013; Litchfield & Donovan, 2016). Medical image perception is a domain of visual expertise that has long recognised the importance of processing the initial glimpse of an image (e.g., Kundel & Nodine, 1975), and given the rapid developments in scene perception research, it is fundamental that these respective research fields are reconciled as both inform each other (Donovan & Litchfield, 2013; Drew, Evans Võ, Jacobson, & Wolfe, 2013a).

A hallmark of expertise is that experts make better and faster decisions than novices (Chase & Simon, 1973; de Groot, 1946/1965; Gobet, 2015) and experts in medical image perception are no exception (for an excellent review of eye movements and visual expertise in medicine see Reingold & Sheridan, 2011; see also this special issue: Fox & Faulkner-Jones, 2017; Szulewski, Kelton, & Howes, 2017). Moreover, examining the expertise-related differences in eye movement patterns reveals the types of visual processing and cognitive strategies that may underlie such expert performance. As mentioned above, our visual system integrates top-down information (e.g., knowledge, expectations) with bottom-up information (visual processing of the incoming image) to make sense of what we see and where to look. Accordingly, experts in a specific domain can draw on their acquired knowledge to be more selective in the information they use to make decisions and are less likely to look at conspicuous but non-informative areas of a scene (for a recent meta-analysis, see Gegenfurtner, Lehtinan, & Säljö, 2011). Indeed, experts in medical image perception are much more selective than novices in deciding where to look (Donovan & Litchfield, 2013; Krupinski, 1996; Kundel & La Follette, 1972; Manning, Ethell, & Crawford, 2003; Manning, Ethell, & Donovan, 2004; Manning, Ethell, Donovan, & Crawford, 2006), and exhibit efficient scanpaths that allow abnormalities to be detected quickly (Krupinski, 1996; Kundel, Nodine, Conant, & Weinstein, 2007; Kundel, Nodine, Krupinski, & Mello-Thoms, 2008), whilst minimizing unnecessary fixations (Manning et al., 2006). Consequently, this allows experts to inspect medical images faster than less experienced observers without compromising on decision accuracy (Nodine, Mello-Thoms, Kundel, & Weinstein, 2002), and there is much to learn from viewing how experts search these complex images (Litchfield, Ball, Donovan, Manning, & Crawford, 2010).

These expert/novice differences can be interpreted by the prominent global-focal search model (Nodine & Kundel, 1987) or its recent formulation, the "holistic model" (Kundel et al., 2007; see also the two-stage detection model by Swensson, 1980). The holistic model proposes that within the first glimpse,

expert observers globally processes the medical image and subsequently make efficient search-related eye movements to potentially abnormal areas to support diagnostic decision-making. This means that prior to foveal search, experts process the low-level information relating to the present image and compare this with their extensive experience of previously viewed normal and abnormal medical images. By drawing on their knowledge of domain-specific visual representations (schema), experts rapidly recognise and coordinate their search for abnormalities and deploy search strategies based on the global information encapsulated within the initial "gist" of image viewing. If nothing is perturbed from the initial global impression, then subsequent search and discovery processing is engaged, which is reliant on individual feature search.

There is substantial converging evidence to support the holistic model, and specifically that expert observers can rapidly process the initial glimpse of the image. The link between globally processing the initial glimpse and diagnostic performance was established by Kundel and Nodine's (1975) tachistoscopic experiments, in which they found that even when images were presented for just 200ms, experts could still correctly detect 70% of abnormal images compared to 97% with no time constraints (see also Carmody, Nodine & Kundel, 1981; Evans, Georgian-Smith, Tambouret, Birdwell, & Wolfe, 2013; Mugglestone, Gale, Cowley & Wilson, 1995; Oestmann et al., 1988). Since these 'flash' studies presented images so quickly that they prevented eye movements and yet performance was still above chance, this provided strong evidence that rapid processing of medical images must be contributing to diagnostic performance, aside from what can be gained by subsequent search and discovery processing. Moreover this led Kundel and Nodine to argue that "visual search begins with a global response that establishes content, detects gross deviations from normal, and organizes subsequent foveal checking fixations" (Kundel & Nodine, 1975, p. 527), and this hypothesis has dominated medical image perception research in the decades since its inception. Indeed, Kundel et al. (2007) and Nodine & Mello-Thoms (2010) proposed that the ability to optimise the processing of the initial glimpse is a hallmark of expertise as they suggested that visual search in medical image perception begins as: SEARCH & DETECT–RECOGNIZE–DECIDE but with experience, this develops into: RECOGNIZE & DETECT–SEARCH–DECIDE.

However, it is important to note that in several of these 'flash' studies, such above-chance performance was only observed when they contained highly conspicuous nodules as performance was much worse for subtle nodules. For example, Carmody et al. (1981) reported that detection of abnormal images did not improve beyond 180ms, but whereas abnormal images containing high and medium visibility nodules were detected often (100% and 83% respectively), abnormal images containing low visibility nodules were only correctly detected 53% of the time. Similarly, Oestmann et al. (1988) found that images containing subtle or obvious cancers were correctly detected 30% and 70% respectively when shown for just 250ms, whereas with unlimited viewing time detection increased to 74% and 98%, respectively. Indeed, in free search conditions Carmody et al., (1981) established that there was a relationship between nodule visibility and the frequency of comparative scans: the poorer the visibility of the nodule, the more likely that eye movements would be directed alternately to normal and abnormal regions to help distinguish pathology from normality. These findings are reflected in the holistic model, in that if the observer does not recognise perturbations in the image then search and discovery processing is undertaken to support diagnostic decision-making. But if the strength of the target signal influences detection so much, how do we know that global processing is involved and that it is not just rapid serial feature search instead? The answer lies with experiments that deliberately disrupted global or 'holistic' processing.

For example, presenting segmented images (Carmody, Nodine, & Kundel, 1980) or rotated images (Oestmann, Greene, Bourgouin, Linetsky, & Llewellyn, 1993) impaired diagnostic performance compared to 'global search', where images were presented normally. In addition, these impairments in detection were independent of target conspicuity as they were found in images containing either obvious or subtle cancers (Oestmann et al., 1993). What is problematic, however, is that since these disruption studies and the 'flash' studies had very limited sample size (3 to 4 observers), and predominately only tested experienced observers (common problems with many medical imaging studies), they actually provide very little direct evidence that experts are the only ones reliant on global processing, or whether in fact this disruption in global processing also impairs less experienced observers. This is an important distinction to make in understanding the

development of visual expertise. For example, it is one thing to infer that experts are able to globally process a domain-specific image and that this helps explain their expert performance, it is another to acknowledge that all observers potentially have access to global processing at some point (for example, in scene perception), and it is just that experts may have fine-tuned an existing system to a fit a particular domain, rather than created a new system entirely from scratch. This has both theoretical and practical implications as to how global processing is used, modified, and if vital to diagnostic performance, how it may be selectively trained to enhance performance. That said, a broader overview of the visual expertise literature shows there are multiple domains where increasing expertise enables observers to shift from local feature processing to holistic processing and thereby take into account the overall configuration of incoming information (Gauthier, Tarr, & Bub, 2010), whether this is expertise in processing particular objects such as faces (Gauthier & Tarr, 1997) dogs (Diamond & Carey, 1986), cars (Curby, Glazek & Gauthier, 2009), fingerprints (Busey & Vanderkolk, 2005), or potentially whole scenes (Kelley, Chun, & Chua, 2003; Werner & Thies, 2000).

Aside from these disruption and flash studies, expert/novice eye movement studies in medical perception have been broadly supportive of the holistic model. A recurring finding of expert/novice studies is that not only can experts correctly identify more abnormalities than novices, the time taken to first fixate abnormalities (search latency) is faster for experts compared to novices (Donovan & Litchfield, 2013; Krupinski, 1996; Kundel, et al., 2007; Kundel, Nodine, Krupinski, & Mello-Thoms 2008; Nodine & Mello-Thoms, 2010; Reingold & Sheridan, 2011). Related to this, experts tend to make longer saccades than novices and these larger eye movements across the image help experts reach targets faster (Krupinski, 1996; Kundel, et al., 2007; Kundel et al., 2008; Manning et al., 2006). According to the holistic model the initial global analysis of the image is thought to initiate and guide search and so this efficiency in expert search behaviour is attributed to experts exploiting global processing that less experienced observers cannot (Kundel et al., 2007). Although low sample size is often a methodological constraint of these expertise studies, efforts have been made to combine time-to-first fixation data from a number of small sample mammography studies and this shows that faster search times are associated with expert performance (Kundel et al., 2008). Kundel et al. (2008) performed a mixture distribution analysis and found that over half of all cancers in mammography were fixated within 1 second, with the remaining cancers fixated in subsequent search. These two distributions of time-to-first fixations were taken as evidence to support the two information processing systems involved in the detection of abnormalities; 1) rapid initial holistic processing, and 2) the slower processing relating to search and discovery. However, as mentioned above, target conspicuity can drive early detection rates and therefore it does not appear to have been considered that instead of these distributions reflecting the two information processes in question, these distributions may simply be the time taken to find obvious and subtle cancers respectively. Indeed, one of the actual studies used by Kundel et al. (2008) found that time to first fixate cancers is dependent on the subtlety of the targets, with subtle cancers taking longer to be fixated (Krupinski, 1995). In addition, we recently showed that unlike mammography, with chest x-rays only 33% of cancers were fixated within 1 second, whereas 56% of cancers were fixated within 2 seconds (Donovan & Litchfield, 2013). In both cases observers would have globally processed the image, but again, it is not clear whether the subsequent differences in search latencies arise because of the differences in how each image modality was globally processed, or simply because our chest x-ray images contained more subtle abnormalities which led to longer search times.

As we have recently argued (Donovan & Litchfield, 2013; Litchfield & Donovan, 2016), the underlying methodological problem with using time-to-first fixation data is that it is obtained from eye tracking experiments under free viewing conditions, whereby the observer has constant access to the whole scene via peripheral vision. This makes it difficult to isolate the specific contribution of processing the initial glimpse of the scene on subsequent eye movement behaviour. Scene perception research suggests that the initial representation evolves during scene viewing, and may provide a frame on which subsequent information can be added (Friedman, 1979). Visual and semantic information obtained from successive fixations on objects and other aspects of the scene can exert influence on future eye movements (Henderson, Weeks, & Hollingworth, 1999). This poses a problem when it comes to examining eye movements during free viewing as it can be difficult to discriminate how eye guidance is affected by the initial scene

representation compared with this continuously updated representation. Kundel et al. (2008) acknowledged that input is continuously gathered from the periphery during search, but are unable to specify precisely how the initial global processing interacts with the ongoing representation and how such information is integrated to guide search. As it stands, our understanding of visual expertise in medical image perception has been derived from different methodological techniques where eye movements were prevented in tachistoscopic 'flash' studies (Kundel et al., 1975; Carmody et al., 1981; Oestmann et al., 1988), or by analysing time-to-first fixation data in which the initial global impression is never dissociated from the ongoing scene representation (Kundel et al., 2008). The recent meta-analysis by Gegenfurtner et al. (2011) confirmed that across a range of visual domains, experts do indeed have shorter time-to-first fixations and make longer saccades compared to non-experts. But whilst this provides even more compelling evidence that experts have faster search latencies and can process domain-specific scenes better than novices, it is still unclear whether it is an expertise-specific advantage in global processing that specifically contributes to search guidance.

Moreover, recent research has shown that global processing may be exploited to rapidly categorize the image, but this does not mean such processing directly supports object recognition. Evans et al. (2013) adopted the traditional 'flash' methodology and required experts and non-experts to rate whether flashed images were abnormal or not under different flash durations. In addition, a subset of observers also had to localize where they thought the abnormality was using a blank outline of the image. By dissociating detection from localization decisions, Evans et al. found that experts could exploit this initial glimpse of the image better than non-experts regardless of the flash duration. Critically however, all groups of observers were only at chance when it came to actually locating the abnormalities. Although Evans et al. did not measure eye movements, these findings are consistent with recent research on scene perception in that global processing appears to enable the gist of the scene to be extracted, but this is just part of a larger system of eye guidance (see Tatler, 2009), and so it is not necessarily the case that global processing itself constrains subsequent search. Given the aforementioned problems with flash and free viewing eye-tracking methodologies, we argued that a new methodology was needed that can dissociate initial scene processing from subsequent search and decision making (Donovan & Litchfield, 2013). We therefore proposed that the recently developed FPMW paradigm (Castelhano & Henderson, 2007) would be a suitable methodology for providing direct evidence of how the initial scene representation may guide search and decision-making and help isolate the contribution of domain-specific visual expertise, compared to our shared expertise at processing real world images in scene perception.

2.   Flash-Preview Moving Window

The FPMW paradigm (Castelhano & Henderson, 2007) draws on the 'flash' methodology that was previously discussed, but also the 'moving window' paradigm. The gaze-contingent 'moving window' paradigm was originally developed by McConkie and Rayner (1975) to investigate the visual span in reading. The 'moving window' allows the observer to continue making eye movements whilst information presented at fovea, parafovea and the periphery is systematically controlled by the researcher and is one of the most powerful eye-tracking methods we have available (Rayner, 2009). This technique alters how much of the scene can be processed by overlaying a variable size mask that is tied to the central fixation position recorded from an eye-tracker and occludes the rest of the scene outside the gaze-contingent moving window. Typically, the observer can examine a scene using their high-resolution fovea but are not able to use parafoveal and peripheral processing as all visual information outside the window is degraded, or simply turned into irrelevant material. Much to the ire of participants, this can also be reversed so that the mask occludes the fovea, and so instead the scene can only be processed by parafoveal and peripheral vision. By changing the size of the window to the point that performance is significantly worse than no window conditions, researchers can systematically assess how much information can be detected and extensively processed outside the high-resolution fovea. Using this paradigm it has been established that experience in reading a particular language changes the size and shape of the visual span (Pollatsek, Bolozky, Well, & Rayner, 1981), and that the visual span increases with expertise, whether this is in terms of reading skill

(Rayner, Slattery, & Bélanger, 2010), or in other domains of visual expertise, such as chess (Charness, Reingold, Pomplun, & Stampe, 2001; Reingold, Charness, Pomplun, & Stampe, 2001; Reingold & Sheridan, 2011).

In the FPMW paradigm (Castelhano & Henderson, 2007), observers are briefly shown a preview of the upcoming search scene and are subsequently asked to search for a particular target object whilst their peripheral vision is restricted to a gaze-contingent moving window. Observers therefore have to rely on what they could process from the initial glimpse of the scene and use this initial representation to guide their search, alongside any other knowledge they have about the target and scene-type (e.g., what the target typically looks like and the likely location of the target given the scene-context). This method offers unique insights as to how the initial glimpse guides eye movement behaviour and draws on the strengths of the previously discussed methodologies. With FPMW, not only is it possible to systematically control what is presented in the scene preview and allow subsequent search-related eye movements to be made, the type of visual processing (e.g., foveal, parafoveal, peripheral) can also be systematically controlled during these eye movements. In the standard setup (see Figure 1a for an example), observers begin by fixating a central fixation cross and then a preview of a scene or control image (for example a visual mask, or different scene type) is briefly flashed (typically for around 250ms), this is then replaced by a visual mask (for 50ms). A word then appears (for around 1000ms-2000ms) indicating which target object must be found. Search then commences and the observer has 15 seconds to find the target whilst their peripheral vision is restricted to a small gaze-contingent moving window (between 2° and 5°). Once the observer has found the target, they press a button whilst looking directly at the target (a gamepad is used to avoid latency delays from keyboards). Subsequent analysis is used to identify whether the target object was correctly identified. This can be done by creating areas of interest (AOI) around the target object and establishing whether or not the observers gaze was within the specified AOI when the button was pressed, with analysis usually restricted to those trials where targets were correctly identified. As this is a gaze-contingent display method, it is important that the display screen is updated with as minimal delay as possible. As such, the FPMW should be presented on a monitor with a high refresh rate (>100hz, preferably CRT to minimise delay) using a desktop eye-tracker with a high sampling rate (>500hz) and a chin rest should be used to maintain accuracy (calibrations should be accepted only if average visual angle error < 0.5°).
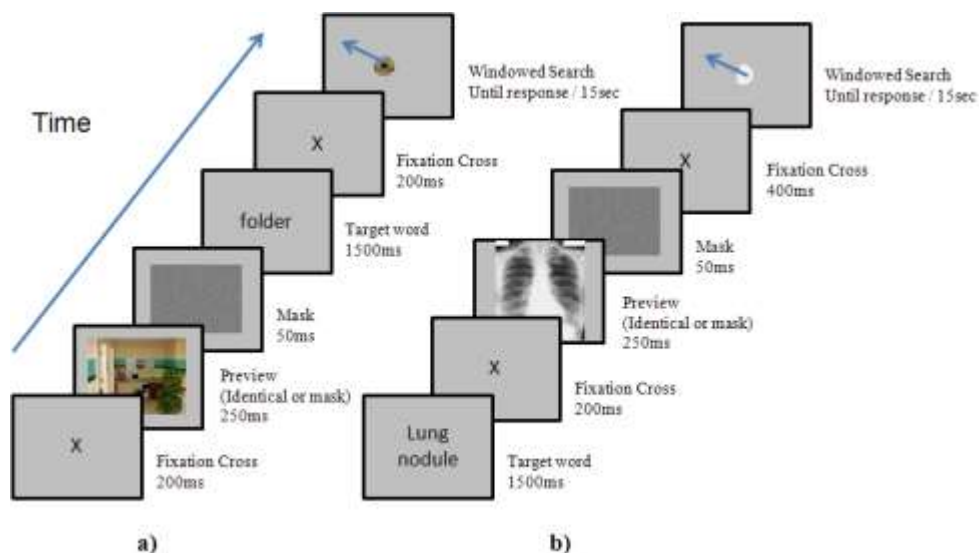

Figure 1. Trial sequence of the 'Flash-Preview Moving Window' paradigm a) the traditional FPMW sequence where target identity is known only after the preview has been encoded b) the adapted trial sequence used by Litchfield and Donovan (2016), where target identity is known before the preview has been encoded.

There are a number of key parameters that can be manipulated in the FPMW depending on the research question (see section 3 below). Specific eye movement metrics and performance measures can be obtained to establish the effects of the initial scene preview on subsequent search and decision-making (see Table 1). Several studies have now used FPMW to investigate scene perception and they all largely demonstrate scene preview benefits in search, whereby the time-to-first fixate targets (search latency) is faster with scene previews compared to control conditions, and other eye movement metrics reflect greater efficiencies in how windowed search is initiated and executed. Note that time-to-first fixation is one of the primary metrics to establish how windowed search is guided by the initial glimpse of a scene using FPMW, however, this metric is distinct from time-to-first fixation used in free viewing conditions, as the latter do not control the level of parafoveal and peripheral processing during search. The FPMW paradigm (Castelhano & Henderson, 2007) has shown that scene preview benefits are obtained when the preview is identical to the subsequent search scene, and this benefit is maintained even if the preview is different in size. However, there is no benefit if the preview image is different to the search scene but belongs to the same scene category. Moreover, the scene preview benefit exists even if the target object was not visible during the preview (i.e., digitally removed), but only found through windowed search, thereby confirming the benefit of scene-context processing, irrespective of any additional local target processing that could occur when targets are present in previews (Castelhano & Henderson, 2007; Võ & Henderson, 2010). The FPMW has also been used to demonstrate how semantically consistent and inconsistent objects are processed within scenes (Castelhano & Heaven 2011; Võ & Henderson, 2011), and how learned object function may guide attention aside from object features (Castelhano & Witherspoon, 2016). The ability to process the scene preview has been linked to individual differences in visual perceptual processing speed (Võ & Schneider, 2010), and the time-course of the initial representation derived from the scene preview has also been investigated. Although 250ms is the standard preview duration, previews as short as 75ms can lead to search advantages compared to no preview conditions, and even as low as 50ms, provided the integration time between scene preview, target word and search was extended from 500ms to 3000ms (Võ & Henderson, 2010). Moreover, preview benefits in search diminish with time, and in fact, the benefit of an initial glimpse of a scene facilitates search only up to the fourth fixation (Hillstrom, Scholey, Liversedge, & Benson, 2012). Indeed, based on these latter FPMW findings we argued that the benefit of the initial glimpse in medical image perception would depend on the difficulty of finding abnormalities within this short time-frame (Donovan & Litchfield, 2013). Taken together, this growing body of research using FPMW provides many useful insights into our shared expertise in scene perception and how we rapidly guide our eye movements based on just a single glimpse of the upcoming scene.

To establish the contribution of domain-specific expertise in eye guidance, we recently applied the FPMW to medical image perception for the first time and examined whether experts are able to guide their eye movement behaviour more effectively than novices, solely upon seeing an initial glimpse (Litchfield & Donovan, 2016). In all experiments the scene preview was either an identical preview of the upcoming scene or a mask preview (i.e., random noise meaning no preview of the upcoming scene). We first compared performance and eye movement behaviour in a standard scene perception task that required objects to be found from real-world scenes. Consistent with previous FPMW studies, both expertise groups showed identical scene preview benefits, in that their eye movements were more efficient at finding the targets in the scene preview condition compared to the mask preview condition. This is consistent with previous research showing that experts superior visual perceptual processing is domain-specific and does not transfer to domain general tasks that draw on shared expertise (Nodine & Krupinski, 1998; Sowden, Davies, & Roling, 2000). The key comparison study, however, was how well these same expert and novice observers were able to find lung nodules (cancer) from chest x-rays in a subsequent FPMW experiment. We found the typical expertise effect in diagnostic performance, with experts being able to identify more cancers than novices. Expert diagnostic performance was the same in scene preview as mask preview. In contrast, novice observers were actually worse at making diagnostic decisions when presented a scene preview prior to search. We interpreted these negative effects of preview on novice performance as being due to interference from distractors, in that by previewing the upcoming medical image novices were exposed to the potential 'nodule-like' distractors that are inherent in chest x-rays, but are in fact normal features. The fact that

expert's decisions were not similarly impaired in scene preview suggested that their more elaborate knowledge of what nodules are ensured they were less biased by these 'normal' distractors.

Table 1.

Key dependent variables obtained using the 'Flash-preview Moving Window' Paradigm.

Dependent Variable	Definition
Performance measures:	
Accuracy (%)	The % of targets correctly identified during search.
Reaction Times (RT)	The time taken from onset of windowed search until search is terminated via button press.
Search-related measures:	
Time-to-first fixation (Search Latency)	This is the total amount of time spent from the onset of the search scene up to (but not including) the first fixation on the target. This is considered one of the primary measures to establish that a scene preview benefit has been found in relation to search.
Number of Fixations	This is the total number of fixations made from the onset of the search scene up to (but not including) the first fixation on the target. Although this measure is related to latency, it can help show whether observers are effectively selecting target candidates for fixation.
Scan Path Ratio.	This is how much of the scene was explored. The scan path ratio is the length of the scan pattern through the scene until the first fixation on the target (total distance between all fixations from scene onset to the first fixation on target) divided by the most direct path to the object (distance from the central fixation point to the centre of the target object). As the ratio approaches 1 this indicates an optimal path to target location (Henderson et al., 1999).
First eye movement of search:	
Initial Saccadic Latency	This is the latency of the first eye movement of search.
Initial Saccadic Amplitude.	This is the amplitude of the first eye movement of search.
Target processing measures:	
First Fixation Duration	This is the duration of the initial fixation on the target and reflects the initial processing of the target.
First Gaze Duration	The total duration of all fixations on the target since it was first fixated until the gaze moves away from the target.
Total Time on Target	This is the summation of all fixation durations on the target before a response button is pressed (including any refixations).

Given the holistic model (Kundel et al., 2007) suggests that experts should be better at searching for targets in these domain-specific scenes, we were surprised to find that there was only a weak scene preview effect in this medical perception task. More specifically, both novices and experts showed a slight search improvement with scene preview, with nodules fixated in fewer fixations and a borderline search latency improvement, however, with each search measure there was no expertise x preview interaction. When we unpacked the effect of scene preview we found that experts demonstrated more of a search benefit (-765 ms) than novices (-260 ms), but the scale of these scene preview search advantages were dwarfed in comparison to the scene preview benefits these same experts and novices demonstrated when searching for targets in scene perception (–1,282ms and –1,620ms respectively). Medical abnormalities are more difficult to identify than targets used in scene perception research and so having such difficult targets to find could have contributed to this weaker scene preview search benefit. However, scene preview effects are still found even when the target is not even visible during scene previews (Castelhano & Henderson, 2007), and so it is likely that the learned spatial associations between target and scene are greatly contributing to the scene preview benefit and where best to look. If the targets in scene perception tasks have a more predictable location (a closer target-scene pairing) then previewing the scene-context would enable those learned target-scene spatial associations to be incorporated and guide eye movements. However, in our medical image perception task we highlighted that nodules may not have such a close target-scene spatial association (Båth et al., 2005), and this may be why we observed such a weak scene preview effect in this medical image perception task. The broader implication to visual expertise is that experts would have learned these spatial associations whereas novices will have not – so whilst we only found weak expertise effects in our specific medical imaging study, it follows that much stronger scene preview effects should be found in tasks where there is a clearer target-scene spatial association for experts to learn, and subsequently exploit. Whilst these issues are not well specified in the holistic model at all, they are clarified in greater detail in the increasing number of eye guidance models in scene perception (for an overview, see Tatler, 2009). Indeed, this once again highlights the growing need to reconcile visual expertise research in medical image perception with current work in scene perception.

Applying FPMW to medical image perception also raised a key methodological distinction between scene perception research and medical image perception tasks. In medical image perception observers are typically searching for the same type of target (e.g., nodule) within the same type of scene (e.g., chest x-ray), whereas in scene perception research the target and scene type are often randomised on each trial. As such, the latter situation may be maximising the advantage of seeing the scene preview relative to mask preview. To rule out the possibility that our weak scene preview effects were simply due to observers repeatedly searching for the same type of target from the same type of scene, a third FPMW experiment required novice and experienced observers to search through a random presentation of three different types of medical images (chest x-rays, brain images, skeletal images), each with their own specific target abnormalities (lung nodules, brain tumors, bone fractures). The rationale was that this should give experienced observers the opportunity to exploit their rapid understanding of these scenes compared to novices and demonstrate a stronger scene preview benefit with increasing expertise. We found that experienced observers were better overall than novices at identifying abnormalities, and their reaction times were faster than novices regardless of the preview condition. Once again novices were worse at identifying targets if given a preview, but controversially, now even experienced observers were impaired at identifying target abnormalities if shown a scene preview before commencing search. Furthermore, we observed no scene preview benefit in search whatsoever. Although the impairments in accuracy for scene preview were relatively small (~3% drop in accuracy), the fact that we observed any such impairment in experienced observers seeing an initial glimpse of the scene goes against our conventional wisdom that processing the initial glimpse of the scene is beneficial to performance. To interpret these counter-intuitive findings and the fact that expert observers were not likewise impaired in Experiment 2, we indicated that targets in medical image perception are difficult to identify even if directly fixating them (Kundel et al., 1978; Donovan & Litchfield, 2013), and that by repeatedly searching the same type of scene expert observers may be able to attenuate the distractors that share similar features with pathology (cf. Kompaniez-Dunigan, Abbey, Boone, & Webster, 2015). Clearly, however, much more research is required to uncover the reasons behind these effects and to better

understand what processes govern whether an initial glimpse does, or does not, lead to more effective search and decision-making.

The purpose of this paper was to highlight the current theoretical and methodological frontier and introduce the FPMW paradigm as a method for investigating visual expertise. However, this is not the first time that gaze-contingent eye-tracking methodologies have been used in medical image perception. Not only were Kundel and Nodine early adopters of eye-tracking research (Kundel & La Follette, 1972; Kundel, Nodine, & Carmody, 1978), they also undertook 'moving window' studies of their own in a bid to dissociate the role central and peripheral vision has on search and decision-making (Kundel, Nodine, & Toto, 1984; Kundel, Nodine, & Toto, 1991). Although data was only based on 2 (Kundel et al., 1984) or 4 expert observers (Kundel et al., 1991), in their study a single chest x-ray was fully viewable at all times and lung nodules that were artificially added to different areas of the image were only visible when they fell within the specified moving window (either 1.5°, 3.5°, 5.5°, 8.5°, or no window). A single chest x-ray was selected for all trials so that the global features of the image were controlled and that the specific role of foveal, parafoveal and peripheral vision could be determined (Kundel et al., 1984). Even from these early studies, it became clear that reducing the size of the moving window to 1.5° dramatically impaired the detection of the nodules and that the time to first fixate nodules decreased with larger windows, consistent with what would later be established more clearly in chess expertise (Reingold et al., 2001) and reading expertise (Rayner et al., 2010). We therefore see the FPMW paradigm as a continuation of this long tradition to understand visual expertise using eye-tracking methodology alongside traditional 'flash' methodologies. Although this FPMW methodology currently focuses on static image interpretation, and not on how dynamic medical images are interpreted (Bertram, Helle, Kaakinen, & Svedström, 2013; Drew et al., 2013b; Phillips et al., 2013) we are nevertheless excited to see how the FPMW will force us to re-evaluate what we think we know about visual expertise in medical image perception. Unlike many of the other methodologies that can be used to investigate visual expertise and are discussed in this special issue, research has only just begun on using FPMW to understand visual expertise. As such, for instructive purposes the remainder of the paper provides an overview of the key parameters that can be manipulated in FPMW.

3. Overview of key FPMW parameters

3.1 Flash Preview (content)

A key strength of the FPMW paradigm is that what is presented in the preview does not have to correspond with what is subsequently searched for using the moving window. The baseline scene preview condition should be identical to the subsequent search scene, but otherwise, performance can be compared to a variety of control preview manipulations (including whether targets are even present or changed in previews).

3.2 Flash Preview Duration

This is how long the preview is presented for and the typical duration is 250ms. Shorter durations can be used to obtain a scene preview benefit (see Võ, & Henderson, 2010), but if longer durations are used eye movements are likely to be made actually within the preview.

3.3 Size of Moving Window

This refers to the size of the moving window. The typical window size is between 2-5 degrees and reducing the size of window is likely to increase search times (cf. Kundel et al., 1984). It is unknown at this time whether there is a relationship between flash preview duration and the size of the moving window.

3.4 Sequence of Preview

In the majority of FPMW studies to date, the observer is first given a scene preview and then afterwards informed what target should be detected in search (Figure 1a). This setup is useful to understand how newly activated target knowledge can be integrated with the currently held visual representation of the scene. However, there are many situations where observers already know the target item they are looking for *before* they actually see the critical search scene. Indeed, in all of the medical image perception studies discussed, observers knew beforehand what was the target (e.g., cancer) before they actually saw the medical image. Accordingly, Litchfield and Donovan (2016) created a modified FPMW sequence where target knowledge was activated before the preview was presented (Figure 1b). Note that robust scene preview benefits were found for scene perception using this modified sequence, even if the same was not the case when medical images were used.

3.5 Integration Time

This refers to the time interval between the scene preview, the target knowledge and the onset of windowed search. Integration can influence the extent to which the scene preview can be exploited to guide search (Võ & Henderson, 2010), and will also be affected by the sequence of preview.

Keypoints

- The 'flash-preview moving window' (FPMW) paradigm (Castelhano & Henderson, 2007) is a new gaze-contingent eye tracking technique that isolates the specific contribution a brief glimpse of a scene has on subsequent eye movement behaviour and decision-making.

- Prevailing theories of medical image perception such as the Holistic model (Kundel et al., 2007) propose that experts exploit global processing of the initial glimpse of medical images to detect abnormalities and guide their eye movements better than novices.

- The FPMW has recently been used for the first time to directly assess the contribution of the initial glimpse on subsequent search behaviour as a function of expertise (Litchfield & Donovan, 2016), but found only weak expertise-specific advantages of processing scene previews.

- Applying the FPMW to a medical image perception task demonstrated that novice observers and (in some situations) experienced observers, were worse at identifying targets when given a scene preview of the upcoming search scene.

- Further research using FPMW in different visual expertise domains is required to build on these preliminary findings.
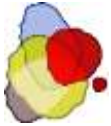
References

Båth, M., Håkansson, M., Börjesson, S., Kheddache, S., Grahn, A., Ruschin, M., . . . Månsson, L. G. (2005). Nodule detection in digital chest radiography: Introduction to the RADIUS chest trial. *Radiation Protection Dosimetry, 114*, 85–91. https://doi.org/10.1093/rpd/nch575

Bertram, R., Helle, L., Kaakinen, J. K., & Svedström, E. (2013). The effect of expertise on eye movement behaviour in medical image perception. *PloS One, 8*, e66169. https://doi.org/10.1371/journal.pone.0066169

Busey, T. A., & Vanderkolk, J. R. (2005). Behavioral and electrophysiological evidence for configural processing in fingerprint experts. *Vision Research, 45*, 431-448. https://doi.org/10.1016/j.visres.2004.08.021

Buswell, G. (1935). *How People Look at Pictures.* University of Chicago Press.

Carmody, D. P., Nodine, C. F., & Kundel, H. L. (1980). Global and segmented search for lung nodules of different edge gradients. *Investigative Radiology, 15*, 224–233.

Carmody, D. P., Nodine, C. F., & Kundel, H. L. (1981). Finding lung nodules with and without comparative visual search. *Perception & Psychophysics, 29*, 594–598. https://doi.org/10.3758/BF03207377

Castelhano, M. S., & Heaven, C. (2011). Scene context influences without scene gist: Eye movements guided by spatial associations in visual search. *Psychonomic Bulletin & Review, 18*, 890-896. https://doi.org/10.3758/s13423-011-0107-8

Castelhano, M. S., & Henderson, J. M. (2007). Initial scene representations facilitate eye movement guidance in visual search. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 753-763. https://doi.org/10.1037/0096-1523.33.4.753

Castelhano, M. S., & Witherspoon, R. L. (2016). How you use it matters: Object function guides attention during visual search in scenes. *Psychological Science, 27*, 606-621. https://doi.org/10.1177/0956797616629130

Charness, N., Reingold, E. M., Pomplun, M., & Stampe, D. M. (2001). The perceptual aspect of skilled performance in chess: Evidence from eye movements. *Memory & Cognition, 29*, 1146-1152. https://doi.org/10.3758/BF03206384

Cohen, M. A., Dennett, D. C., & Kanwisher, N. (2016). What is the bandwidth of perceptual experience? *Trends in Cognitive Sciences, 20*, 324-335. https://doi.org/10.1016/j.tics.2016.03.006

Curby, K. M., Glazek, K., & Gauthier, I. (2009). A visual short-term memory advantage for objects of expertise. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 94-107. https://doi.org/10.1037/0096-1523.35.1.94

de Groot, A. D. (1946). *Het denken van den schaker*. Amsterdam: Noord Hollandsche.

de Groot, A. D. (1965). *Thought and choice in chess*. The Hague, Netherlands: Mouton.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: an effect of expertise. *Journal of Experimental Psychology: General, 115*, 107-117. https://doi.org/10.1037/0096-3445.115.2.107

Donovan, T., & Litchfield, D. (2013). Looking for cancer: Expertise related differences in searching and decision making. *Applied Cognitive Psychology, 27*, 43-49. https://doi.org/10.1002/acp.2869

Drew, T., Evans, K. K., Võ, M. L. H., Jacobson, F. L., & Wolfe, J. M. (2013a). Informatics in radiology: What can you see in a single glance and how might this guide visual search in medical images? *Radiographics, 33*, 263–274. https://doi.org/10.1148/rg.331125023

Drew, T., Vo, M. L. H., Olwal, A., Jacobson, F., Seltzer, S. E., & Wolfe, J. M. (2013b). Scanners and drillers: Characterizing expert visual search through volumetric images. *Journal of Vision, 13*, 3. https://doi.org/10.1167/13.10.3

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*, 1170-1175. https://doi.org/10.3758/s13423-013-0459-3

Fei-Fei, L., Iyer, A., Koch, C., & Perona, P. (2007). What do we perceive in a glance of a real-world scene? *Journal of Vision, 7*, 1-29. https://doi.org/10.1167/7.1.10

Fox, E. S., & Faulkner-Jones, B. E. (2017). Eye-tracking in the study of visual expertise: Methodology and approaches in medicine. *Frontline Learning Research, 5*, 29-40. https://doi.org/10.14786/flr.v5i3.258.

Friedman A. (1979). Framing pictures: the role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General, 108*, 316-355. https://doi.org/10.1037/0096-3445.108.3.316

Gauthier, I., & Tarr, M. J. (1997). Becoming a "Greeble" expert: Exploring mechanisms for face recognition. *Vision Research, 37*, 1673-1682. https://doi.org/10.1016/S0042-6989(96)00286-6

Gauthier, I., Tarr, M., & Bub, D. (2009). *Perceptual expertise: Bridging brain and behavior*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195309607.001.0001

Gegenfurtner, A., Lehtinen, E., & Säljö, R. (2011). Expertise differences in the comprehension of visualizations: a meta-analysis of eye-tracking research in professional domains. *Educational Psychology Review, 23*, 523-552. https://doi.org/10.1007/s10648-011-9174-7

Gobet, F. (2015). *Understanding expertise: A multidisciplinary approach*. London: Palgrave.

Henderson, J. M. (2007). Regarding scenes. *Current Directions in Psychological Science, 16*, 219-222. https://doi.org/10.1111/j.1467-8721.2007.00507.x

Henderson, J. M., & Hollingworth, A. (1999). The role of fixation position in detecting scene changes across saccades. *Psychological Science, 5*, 438-443. https://doi.org/10.1111/1467-9280.00183

Henderson, J. M., Weeks, P. A., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 210-228. https://doi.org/10.1037/0096-1523.25.1.210

Henderson, J. M., Pollatsek, A., & Rayner, K. (1989). Covert visual attention and extrafoveal information use during object identification. *Perception & Psychophysics, 45*, 196-208. https://doi.org/10.3758/BF03210697

Hillstrom, A. P., Scholey, H., Liversedge, S. P., Benson, V. (2012). The effect of the first glimpse at a scene on eye movements during search. *Psychomic Bulletin & Review, 19*, 204-210. https://doi.org/10.3758/s13423-011-0205-7

Just, M. A., & Carpenter, P. A. (1984). Using eye fixations to study reading comprehension. In D. E. Kieras & M. A. Just (Eds.), *New methods in reading comprehension research* (pp. 151–182). Hillsdale: Erlbaum.

Kelley, T. A., Chun, M. M., & Chua, K. P. (2003). Effects of scene inversion on change detection of targets matched for visual salience. *Journal of Vision, 3*, 1-5. https://doi.org/10.1167/3.1.1

Kompaniez-Dunigan, E., Abbey, C. K., Boone, J. M., & Webster, M. A. (2015). Adaptation and visual search in mammographic images. *Attention, Perception, & Psychophysics, 77*, 1081-1087. https://doi.org/10.3758/s13414-015-0841-5

Krupinski, E. A. (1995). Visual search of mammographic images: influence of lesion subtlety. *Academic Radiology, 12*, 965-969. https://doi.org/10.1016/j.acra.2005.03.071

Krupinski, E. A. (1996). Visual scanning patterns of radiologists searching mammograms. *Academic Radiology, 3*, 137-144. https://doi.org/10.1016/S1076-6332(05)80381-2

Kundel, H. L., & La Follette, P. S. (1972). Visual search patterns and experience with radiological images. *Radiology, 103*, 523–528. https://doi.org/10.1148/103.3.523

Kundel, H. L., & Nodine, C. F. (1975). Interpreting chest radiographs without visual search. *Radiology, 116*, 527–532. https://doi.org/10.1148/116.3.527

Kundel, H.L., Nodine, C.F., & Carmody, D.P. (1978). Visual scanning, pattern recognition, and decision making in pulmonary nodule detection. *Investigative Radiology, 13*, 175–181. https://doi.org/10.1097/00004424-197805000-00001

Kundel, H. L., Nodine, C. F., Conant, E. F., & Weinstein, S. P. (2007). Holistic component of image perception in mammogram interpretation: gaze-tracking study 1. *Radiology, 242*, 396-402. https://doi.org/10.1148/radiol.2422051997

Kundel, H. L., Nodine, C. F., Krupinski, E. A., Mello-Thoms, C. (2008). Using gaze-tracking data and mixture distribution analysis to support a holistic model. *Academic Radiology, 15*, 881–886. https://doi.org/10.1016/j.acra.2008.01.023

Kundel, H. L., Nodine, C. F., & Toto, L. C. (1984). Eye movements and the detection of lung tumors in chest images. In A. G. Gale, & F. Johnson (Eds.). *Theoretical and Applied Aspects of Eye Movement Research* (pp. 297-304). North-Holland: Elsevier.
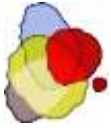
Kundel, H. L., Nodine, C. F., & Toto, L. C. (1991). Searching for lung nodules. The guidance of visual scanning. *Investigative Radiology, 26*, 777-781.

Litchfield, D., Ball, L. J., Donovan, T., Manning, D. J., & Crawford, T. (2010). Viewing another person's eye movements improves identification of pulmonary nodules in chest x-ray inspection. *Journal of Experimental Psychology: Applied, 16*, 251-262. https://doi.org/10.1037/a0020082.

Litchfield, D., & Donovan T. (2016). Worth a quick look? Initial scene previews can guide eye movements as a function of domain-specific expertise but can also have unforeseen costs. *Journal of Experimental Psychology: Human Perception and Performance*, *42*, 982-994. https://doi.org/10.1037/xhp0000202

Manning, D. J., Ethell, S., & Crawford, T. (2003). An eye-tracking AFROC study of the influence of experience and training on chest x-ray interpretation. *SPIE, 5034*, 257–266. https://doi.org/10.1117/12.479985

Manning, D. J., Ethell, S. C., & Donovan, T. (2004). Detection or decision errors? Missed lung cancer from the posteroanterior chest radiograph. *British Journal of Radiology, 77*, 231-235. https://doi.org/10.1259/bjr/28883951

Manning, D. J., Ethell, S. C., Donovan, T., & Crawford, T. J. (2006). How do radiologists do it? The influence of experience and training on searching for chest nodules. *Radiography, 12*, 134-142. https://doi.org/10.1016/j.radi.2005.02.003

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics, 17*, 578-86. https://doi.org/10.3758/BF03203972

Mugglestone, M. D., Gale, A. G., Cowley, H. C., & Wilson, A. R. M. (1995). Diagnostic performance on briefly presented mammographic images. *SPIE, 2436*, 106–115. https://doi.org/10.1117/12.206840

Nodine, C. F., & Krupinski, E. A. (1998). Perceptual skill, radiology expertise, and visual test performance with NINA and WALDO. *Academic radiology, 5*, 603-612. https://doi.org/10.1016/S1076-6332(98)80295-X

Nodine, C. F., & Kundel, H. L. (1987). The cognitive side of visual search in radiology. In J. K., O'Regan, A. Levy-Schoen (Eds). *Eye movements: From physiology to cognition* (pp. 573–582). Amsterdam: Elsevier. https://doi.org/10.1016/B978-0-444-70113-8.50081-3

Nodine, C. F., & Mello-Thoms, C. (2010). The role of expertise in radiologic image interpretation. In E. Samei, E. Krupinski (Eds.), *The Handbook of Medical Image Perception and Techniques* (pp. 139–156), New York: Cambridge University Press.

Nodine, C. F., Mello-Thoms, C., Kundel, H. L., & Weinstein, S. P. (2002). Time course of perception and decision making during mammographic interpretation. *American Journal of Roentgenology, 179*, 917-923. https://doi.org/10.2214/ajr.179.4.1790917

Oestmann, J. W., Greene, R., Bourgouin, P. M., Linetsky, L., & Llewellyn, H. J. (1993). Chest "gestalt" and detectability of lung lesions. *European Journal of Radiology, 16*, 154-157. https://doi.org/10.1016/0720-048X(93)90015-F

Oestmann, J. W., Greene, R., Kushner, D. C., Bourgouin, P. M., Linetsky, L., & Llewellyn, H. J. (1988). Lung lesions: correlation between viewing time and detection. *Radiology, 166*, 451-453. https://doi.org/10.1148/radiology.166.2.3336720

Oliva, A., & Schyns, P. G. (1997). Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive Psychology, 34*, 72-107. https://doi.org/10.1006/cogp.1997.0667

Phillips, P., Boone, D., Mallett, S., Taylor, S. A., Altman, D. G., Manning, D., ... & Halligan, S. (2013). Method for Tracking Eye Gaze during Interpretation of Endoluminal 3D CT Colonography: Technical Description and Proposed Metrics for Analysis. *Radiology, 267*, 924-931. https://doi.org/10.1148/radiol.12120062

Pollatsek, A., Bolozky, S., Well, A. D., & Rayner K. (1981). Asymmetries in the perceptual span for Israeli readers. *Brain and Language, 14*, 174-180. https://doi.org/10.1016/0093-934X(81)90073-0

Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory, 2*, 509-522. https://doi.org/10.1037//0278-7393.2.5.509

Reingold, E. M., & Sheridan, H. (2011). Eye movements and visual expertise in chess and medicine. In S. P. Liversedge, I. D. Gilchrist, S. Everling (Eds). *The Oxford Handbook of Eye Movements* (pp. 767–786). Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199539789.013.0029

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*, 1457-1506. https://doi.org/10.1080/17470210902816461

Rayner, K., & Pollatsek, A. (1981). Eye movement control during reading: Evidence for direct control. *The Quarterly Journal of Experimental Psychology, 33*, 351-373. https://doi.org/10.1080/14640748108400798

Rayner, K., Slattery, T. J., & Bélanger, N. N. (2010). Eye movements, the perceptual span, and reading speed. *Psychonomic Bulletin & Review, 17*, 834-839. https://doi.org/10.3758/PBR.17.6.834

Rayner, K., Smith, T. J., Malcom, G. L., & Henderson, J. M. (2009). Eye movements and visual encoding during scene perception. *Psychological Science, 20*, 6-10. https://doi.org/10.1111/j.1467-9280.2008.02243.x

Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time-and spatial-scale-dependent scene recognition. *Psychological Science, 5*, 195-200. https://doi.org/10.1111/j.1467-9280.1994.tb00500.x

Sowden, P. T., Davies, I. R., & Roling, P. (2000). Perceptual learning of the detection of features in X-ray images: a functional role for improvements in adults' visual sensitivity? *Journal of Experimental Psychology: Human Perception and Performance, 26*, 379–390. https://doi.org/10.1037/0096-1523.26.1.379

Swensson, R. G. (1980). A two-stage detection model applied to skilled visual search by radiologists. *Perception & Psychophysics, 27*, 11-16. https://doi.org/10.3758/BF03199899

Szulewski, A., Kelton, D., & Howes, D. (2017). Pupillometry as a tool to study expertise in medicine. *Frontline Learning Research, 5*, 52-62. https://doi.org/10.14786/flr.v5i3.256

Tatler, B. W. (2009). Current understanding of eye guidance. *Visual Cognition, 17*, 777-789. https://doi.org/10.1080/13506280902869213

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review, 113*, 766-786. https://doi.org/10.1037/0033-295X.113.4.766

Võ, M. L. H., & Henderson, J. M. (2010). The time course of initial scene processing for eye movement guidance in natural scene search. *Journal of Vision, 10*, 1-13. https://doi.org/10.1167/10.3.14

Võ, M. L. H., & Henderson, J. M. (2011). Object-scene inconsistencies do not capture gaze: evidence from the flash-preview moving window paradigm. *Attention, Perception & Psychophysics, 73*, 1742–1753. https://doi.org/10.3758/s13414-011-0150-6

Võ, M. L. H., & Schneider, J. M. (2010). A glimpse is not a glimpse: Differential processing of flashed scene previews leads to differential target search benefits. *Visual Cognition, 18*, 171-200. https://doi.org/10.1080/13506280802547901

Werner, S., & Thies, B. (2000). Is" change blindness" attenuated by domain-specific expertise? An expert-novices comparison of change detection in football images. *Visual Cognition, 7*, 163-173. https://doi.org/10.1080/135062800394748

Wolfe, J. M., Võ, M. L.-H., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes involves selective and non-selective pathways. *Trends in Cognitive Sciences, 15*, 77-84. https://doi.org/10.1016/j.tics.2010.12.001

Yarbus, A. (1967). *Eye movements and vision*. New York: Plenum Press. https://doi.org/10.1007/978-1-4899-5379-7

Zelinsky, G. J., & Schmidt, J. (2009). An effect of referential scene constraint on search implies scene segmentation. *Visual Cognition, 17*, 1004-1028. https://doi.org/10.1080/13506280902764315