Central Lancashire Online Knowledge (CLoK)

Title	Viewing another person's eye movements improves identification of pulmonary nodules in chest x-ray inspection
Type	Article
URL	https://clok.uclan.ac.uk/6118/
DOI	https://doi.org/10.1037/a0020082
Date	2010
Citation	Litchfield, Damien, Ball, Linden, Donovan, Tim, Manning, David J. and Crawford, Trever (2010) Viewing another person's eye movements improves identification of pulmonary nodules in chest x-ray inspection. Journal of Experimental Psychology: Applied,, 16 (3). pp. 251-262. ISSN 1076-898X
Creators	Litchfield, Damien, Ball, Linden, Donovan, Tim, Manning, David J. and Crawford, Trever

Running Head: Improving Chest X-Ray Analysis Using Eye Movement Cues

Viewing Another Person's Eye Movements Improves Identification

of Pulmonary Nodules in Chest X-Ray Inspection

Damien Litchfield[1]*

Linden J. Ball[1]

Tim Donovan[2]

David J. Manning[2]

Trevor Crawford[1]

[1]Department of Psychology,

Lancaster University, UK

[2]School of Medical Imaging Sciences,

Cumbria University, UK

_____

*Corresponding author:

Department of Psychology, Lancaster University, Lancaster, LA1 4YF, UK.

Email: d.litchfield@lancaster.ac.uk

Tel: +44 (0)1524 594742

Fax: +44 (0)1524 593744

Abstract

Double reading of chest x-rays is often used to ensure that fewer abnormalities are missed, but very little is known about how the search behavior of others affects observer performance. A series of experiments investigated whether radiographers benefit from knowing where another person looked for pulmonary nodules, and whether the expertise of the model providing the search behavior was a contributing factor. Experiment 1 compared the diagnostic performance of novice and experienced radiographers examining chest x-rays and found that both groups performed better when shown the search behavior of either a novice radiographer or an expert radiologist. Experiment 2 established that benefits in performance only arose when the eye movements shown were related to the search for nodules; however, only the novices' diagnostic performance consistently improved when shown the expert's search behavior. Experiment 3 re-examined the contribution of task, image, and the expertise of the model underlying this benefit. Consistent with Experiment 1, novice radiographers were better at identifying nodules when shown either a naïve's search behavior or an expert radiologist's search behavior, but they demonstrated no improvement when shown a naïve model not searching for nodules. Our results suggest that although the benefits of this form of attentional guidance may be short-lived, novices can scaffold their decisions based on the search behavior of others.

Viewing Another Person's Eye Movements Improves Identification

of Pulmonary Nodules in Chest X-Ray Inspection

Around 40% of all hospital procedures require some form of medical image interpretation (Beam, Krupinski, Kundel, Sickles, & Wagner, 2006). For successful diagnosis it is essential that those who analyze these images are able to interpret correctly the information that they contain. When examining chest x-rays one of the many features that radiologists look for is the presence of pulmonary nodules because these can be precursors to lung cancer. The radiologist must correctly distinguish these subtle abnormalities from normal features. However, the complexity of medical images of lungs can lead to relatively high numbers of false positive decisions (thinking that a feature is a nodule when, in fact, it is not) and false negative decisions or "misses" (thinking that a feature is normal when, in reality, it is a nodule; see Fitzgerald, 2001; Taylor, 2007).

Previous work on false negative errors suggests that at least 30% arise from insufficient visual search (Kundel, Nodine, & Carmody, 1978; Manning, Ethell, & Donovan, 2004), with novice radiographers being particularly prone to making such search-related errors. With experience, however, decision-making and interpretation processes become the source of most errors (Donovan & Manning, 2006). Decision-making and interpretation skills require appropriate training and extensive exposure to case studies, which makes such skills difficult to develop rapidly. In contrast, viewing the eye movement behavior of another observer completing the same task may help with the development of visual scanning and recognition skills, as this could provide an example of an effective search process (e.g., following where another person looks may reveal to an observer the most task-relevant areas in an image such as probable nodule sites).

It could be argued that experts would be better than novices at guiding observers' attention to relevant task-specific areas, because experts would be more likely to locate

these areas in the first place (e.g., Kundel & Nodine, 1983; Nodine, Mello-Thoms, Kundel, & Weinstein, 2002). However, it is not simply the case that when observers fixate at nodule locations the correct decision is made. Indeed, the more controversial implication of Kundel et al.'s (1978) findings concerning false negative errors is that around 70% of nodules that are missed are in fact fixated (a finding replicated in domains such as baggage screening; McCarley, Kramer, Wickens, Vidoni, & Boot, 2004). As such, these nodule locations would have been given some degree of processing but were deemed to be normal. Making observers re-evaluate their decisions based on these perceptual factors is one way to reduce errors. For example, when radiologists are shown their own eye movement behavior during pulmonary nodule identification, this can lead to an improvement in performance (Krupinski, Nodine, & Kundel, 1993; Kundel, Nodine, & Krupinski, 1990). However, recent evidence using a similar paradigm shows that this perceptual feedback is mostly beneficial for radiographers in the early stages of learning (Donovan, Manning, & Crawford, 2008). The present study aims to expand our understanding of the factors that may improve pulmonary nodule identification by investigating whether viewing *another person's* eye movements improves visual search and decision-making during a nodule detection task.

The Communicative Role of Eye Movements

The eye movements of experts and novices reveal how visual search strategies can change as a function of expertise (Chapman & Underwood, 1998; Rayner, 1998), however, eye movements can also be communicative and direct the attention of others (Kleinke, 1986). Studies have shown that we are highly sensitive to other people's gaze (Gibson & Pick, 1963; Symons, Lee, Cedrone, & Nishimura, 2004) and that this sensitivity develops early, with very young infants demonstrating gaze following behavior (e.g., Brooks & Meltzoff, 2005; Corkum & Moore, 1995). Following another's

gaze can guide attention towards a particular item, which allows for knowledge to be inferred or directly shared if the observers verbalise their interest and intent towards the jointly attended item (Flom & Pick, 2007; Hanna & Brennan, 2007). This "joint attention" has been claimed to underpin much of early learning (Butterworth & Jarret, 1991), but is still highly prevalent in adulthood (Driver et al., 1999; though see Vecera & Rizzo, 2006, for discussion of voluntary aspects of gaze orienting). In particular, adult observers take into account nearby objects when following another's gaze, with recent neuropsychological evidence indicating that following a person's gaze can transfer that individual's intentionality to the observer (Becchio, Bertone, & Castiello, 2008; Lobmaier, Fischer, & Schwaninger, 2006). Thus by watching where another person looks we alter our processing of objects based on the processing of others.

If a person's eye movements can direct our attention and we can also learn to make use of these cues wherever possible, then how much of cognition is shaped this way? Eye movement behavior is often understood as a reflection of underlying cognitive activity (Vivianni, 1990), but recent research suggests that specific eye movement behaviors can also influence cognition (Grant & Spivey, 2003; Pomplun, Ritter, & Velichkovsky, 1996; Thomas & Lleras, 2007). Given the tendency for people to use another's gaze to drive their own reasoning and decision-making, there is currently much interest in presenting the eye movement patterns of one observer to subsequent observers to explore how this form of attentional guidance improves performance in a range of visual search and problem solving tasks. In particular, because experts are known to look at task-relevant areas more often than novices and demonstrate effective search strategies (e.g., Chapman & Underwood, 1998; Charness, Reingold, Pomplun, & Stampe, 2001; Krupinski, 1996), there is a growing belief that an expert's eye movement patterns should be particularly useful in training novices where to look (Krupinski et al., 2006;

Nalanagula, Greenstein, & Gramopadhye, 2006; Sadasivian, Greenstein, Gramopadhye, & Duchowski, 2005).

By recording the eye movement behavior of experts and showing the projected fixation position to other observers, recent studies have demonstrated that novices can detect more faults during aircraft inspection (Sadasivian et al., 2005) and circuitry board inspection (Nalanagula et al., 2006). For tasks such as program debugging and puzzle solving, which involve factors in addition to visual search, viewing another's gaze can result in shorter task completion times (Stein & Brennan, 2004; Velichkovsky, 1995). However, none of these studies systematically examined the role of either model expertise or participant expertise on task performance, such that it remains unclear whether a novice's eye movements could also act as useful attentional cues. In addition, there are some situations where unambiguous verbal comments are preferred to another's gaze (e.g., Bard et al., 2007; Van Gog, Jarodzka, Scheiter, Gerjets, & Paas, 2009). As such, positive effects will likely depend on the task demands and the different ways in which eye movement patterns can be presented (cf. Nalanagula et al., 2006).

A final issue is that projecting gaze behavior to observers involves "artificially" represented gaze (eye-movement patterns dynamically overlaid on a screen-based image of a task), and therefore it is debatable whether the cognitive processes are the same as those during normal gaze perception and gaze following behavior. For example, the emotional expression of the person being watched is often taken into account when using normal gaze as a predictive cue (Bayliss, Frischen, Fenske, & Tipper, 2007) or when determining mutual gaze (Lobmaier, Tiddeman, & Perret, 2008). Although this information is absent when using artificial gaze, this method does provide an opportunity to look at more complex gaze sequences rather than simple directional processes. Indeed, there have been reports of higher order search strategies based on another's real-time

gaze. For example, observers can regulate their own search behavior in a collaborative

visual search task by strategically ignoring areas that they can see are being observed by

their collaborator (Brennan, Chen, Dickinson, Neider, & Zelinsky, 2008). By taking

advantage of these non-verbal gaze cues, observers have been shown to be able to reduce

their search times significantly. Thus, even with artificially represented gaze, observers

alter their information processing behavior based on the processing of others.

Experiment 1

The evidence described above suggests that another's gaze is often used to help

direct attention and facilitate learning. Experiment 1 aimed to examine whether viewing

another's search behavior would improve an observer's ability to identify pulmonary

nodules. However, training requirements will differ depending on an observer's expertise,

such that an observer's experience should be accounted for when adopting new

instructional techniques. It is expected that showing eye movement patterns of an expert

to novice radiographers will lead to improved decision-making performance. Conversely,

showing an expert's eye movement patterns to experienced radiographers may actually

disrupt diagnostic performance (e.g., Donovan et al., 2008; Donovan, Manning, Philips,

Higham, & Crawford, 2005), a possibility that links to the "expertise reversal effect"

(Kalyuga, 2007), whereby instructional techniques designed to reduce cognitive load may

inhibit the application of pre-existing skills. We designed Experiment 1 to examine these

latter predictions relating to observer expertise. We also investigated whether model

expertise had a modulating influence so as to test the assumption that viewing an expert

model leads to greater improvement compared to viewing a novice model.

The experiment involved three main conditions: (1) "free search", where

observers could immediately make decisions regarding nodules; (2) "image preview",

where observers were required to wait for 20s before making a decision; and (3) "eye

movement preview", where observers viewed either an expert's or a novice's eye movements for 20s before making a decision, with half shown a novice's and half shown an expert's eye movements. The free search condition provided the closest approximation to natural performance during nodule detection, whereas the image preview condition controlled for the additional image exposure in the eye movement preview condition.

Method

Participants. The study involved 24 novice radiographers (9 male; 15 female) and 24 experienced radiographers (8 male; 16 female). Participants had normal or corrected-to-normal vision and were paid £6. Novices were all recruited in the final half of their first year radiography degree, which allowed them to accrue several months of formal instruction, and 12 weeks of clinical experience. Hence, novices had at least some experience of medical image perception. Experienced radiographers were enrolled on postgraduate radiography courses and had at least 5 years experience interpreting medical images from a range of modalities, such as ultrasound, computed-tomography, and MRI. The experienced radiographers were expected to be more proficient than novice radiographers in this task, but less skilled than experts who deal primarily with chest x-ray interpretation.

Materials. Seven normal and seven abnormal images were selected for each viewing condition from a test bank of 120 posteroanterior chest x-rays. Each viewing condition contained a total of 16 nodules: 1 image with 5 nodules, 2 images with 3 nodules, 1 image with 2 nodules, and 3 images with 1 nodule. Nodules ranged in size from approximately $0.5°$ to $3°$. Nodules were controlled across conditions for size, shape and conspicuity and were verified by a consultant radiologist. Images were standardized to fit an LCD display (resolution of 1280 x 1024) and were presented using ClearView 2.6. (Tobii Technology, Stockholm, Sweden). Image order was randomised.

Design. The experiment employed a 2 x 2 x 3 mixed design with three factors: participant expertise with two levels (novice vs. experienced), model expertise with two levels (novice vs. expert), and viewing condition with three levels (free search; image preview; eye movement preview). The presentation of the three viewing conditions was completely counterbalanced. In the free search condition participants could make decisions regarding nodules immediately. In the image preview condition participants viewed each x-ray for 20s before being allowed to make a decision.

The eye movement preview condition was identical to the image preview condition except that during the 20s viewing period participants were also shown an overlay of the dynamic eye movements of a pre-selected scanpath (see Nalanagula et al., 2006, for evidence that dynamic and hybrid gaze cues produce enhanced benefits in a visual-search training context when compared with static cues). For each presented scanpath the gaze cursor increased in size in accord with fixation duration. At any one time 500ms of fixation activity was shown, with transitions between fixations depicted by lines (see Figure 1). Participants were told they would be "shown an eye movement recording of someone else examining the image" and that after the recording finished they would be "presented with the same image and asked to search and identify nodules." The only description of the eye movement recordings given to participants was that "the circle represents where someone looked, and the bigger the circle, the longer the individual looked at that particular area." No information was provided regarding the expertise of the model, nor whether the model had provided a correct diagnosis.

<< Insert Figure 1 about here >>

To obtain the model scanpaths, five radiologists (experts) and five first-year radiographers (novices) were paid to complete the nodule detection task under free search conditions using the film test bank described above. Their eye movements were recorded

using a Tobii x50 eye tracker (Tobii Technology, Stockholm, Sweden), a standalone

remote eye-tracking device with an accuracy of 0.5 degrees and a sampling rate of 50Hz.

The minimum fixation duration was set to 100ms with a fixation radius of 50 pixels.

We used the best performing expert and the worst performing novice as the

models to show to the participants in the main study. To assess the overall performance

of each of these five experts and five novices (and the performance of observers in the

main study), a typical approach would be to use signal detection theory, which involves

fitting a standard receiver operating characteristic (ROC) curve based on whether a series

of images are correctly identified as being abnormal or normal (Metz, 2008). This method

overcomes the problem that scoring the percentage of correct decisions neither reflects

decision uncertainty nor the relatively high number of false positives that arise in medical

image perception. However, a fundamental limitation of ROC methodology is that it is

only suitable for binary decisions (e.g., normal/abnormal) and cannot account for the

multiple decisions and/or abnormalities that can occur on the same image (Chakraborty,

2000). Moreover, the requirements of ROC allow an observer to decide an image is

abnormal without actually specifying the location of any abnormalities.

Based on these limitations, the present study adopted a free-response (FROC)

method that required observers to locate and identify any nodules that were present, or

decide that the image was normal. Currently, the most appropriate method of analysing

FROC data is to perform a jackknife free-response ROC (JAFROC) analysis as this has

shown greater statistical power than ROC and "alternative FROC" (AFROC) analysis

(Chakraborty, 2006; Chakraborty & Berbaum, 2004). By applying the non-parametric

Mann-Whitney-Wilcoxon U-statistic to the nodule ratings and the ROC-equivalent

ratings of normal images, a "figure-of-merit" is estimated, which represents the

likelihood that a true positive will be given a higher rating than a false positive. This

single measure of observer performance takes into account continuous ratings of confidence and reflects the error in decision-making in a FROC designed study.

With .50 being at chance level, a high figure-of-merit (e.g., above .75) would mean that the majority of abnormalities were confidently identified, with few detrimental false positive decisions. However, a low figure-of-merit (e.g., below .50) would mean that even though some abnormalities may have been identified, performance would have been offset by a substantial number of detrimental false positive decisions. Accordingly, the eye movement patterns from the novice with the lowest figure-of-merit score (.66) and the radiologist with the highest figure-of-merit score (.79) were selected as the model scanpaths for the eye movement preview condition. Half the participants in the eye movement preview condition were shown the expert model and half the novice model.

Procedure. Participants were seated approximately 60cm from the display and were asked to search for nodules from the test bank described previously. Five practice images were shown before the experiment, demonstrating how nodules differed in size, shape and conspicuity. When participants believed they had identified a nodule they used the mouse to indicate its location and clicked the left mouse button to confirm their decision. Once the decision was made, participants rated how confident they were by calling out a number between 1 and 4, with 4 being "highly confident." Participants pressed a key when they thought they had identified all the nodules on an image, or if they thought an image contained no nodules. Participants were not told how many nodules could be present on each image and were unaware of the prevalence of nodules (50%) across images. The whole experiment took approximately 1 hour to complete.

Results and Discussion

Ratings of true positives and false positives were subjected to a JAFROC analysis as described above (see Chakraborty & Berbaum, 2004). The figure-of-merit scores

obtained from the JAFROC analysis were then examined using a mixed design ANOVA, where the between factors were participant expertise (novice vs. experienced) and model expertise (novice vs. expert), and where the within factor was viewing condition (free search; image preview; eye movement preview). This analysis revealed a main effect of viewing condition [$F(2, 88) = 30.47$, $p < .001$, partial $\eta^2 = .41$], with performance improvements being evident for both novice and experienced radiographers in eye movement preview compared to image preview and free search (see Table 1). We note, however, that the magnitude of the difference was relatively small. Participant expertise was not significant [$F(1, 44) = 2.00$, $p > .16$, partial $\eta^2 = .04$], however, a reliable interaction was found between participant expertise and viewing condition [$F(2, 88) = 4.29$, $p < .017$, partial $\eta^2 = .09$]. The main effect of model expertise was not significant [$F(1, 44) = 3.16$, $p > .08$, partial $\eta^2 = .07$], and model expertise did not interact with any other factors [all Fs < 2.19, all ps $> .12$]. These latter findings suggest that the level of expertise of the model that provided the eye movement recordings was not an influential factor in Experiment 1.

<< Insert Table 1 about here >>

To explore the significant interaction between participant expertise and viewing condition we conducted a series of simple main effects analyses. Both novice and experienced radiographers showed a reliable performance difference across conditions [$F(2, 46) = 25.24$, $p < .001$, partial $\eta^2 = .52$ for novices; $F(2, 46) = 6.84$, $p < .002$, partial $\eta^2 = .23$, for experienced]. Bonferroni-corrected pairwise comparisons confirmed that for both novices and experienced radiographers the eye movement preview condition led to higher scores than the free search and image preview conditions (all ps $< .05$). Simple main effects comparing across participant expertise levels at each condition only revealed a reliable difference for image preview [$F(1, 93.95) = 6.24$, $p < .014$]. These analyses

clarify that both groups benefit from receiving eye movement patterns before making decisions regarding nodules when compared with free search or image preview. The analyses also indicate that the cause of the interaction between participant expertise and viewing condition relates to the performance separation between experienced and novice radiographers in the image preview condition.

The overall decision times for each condition (see Table 1) were analyzed to examine if eye movement preview increased decision times, as seen in studies of aircraft inspection (Sadasivian et al., 2005). An equivalent mixed design ANOVA was employed as in the analysis of JAFROC scores. Decision time was measured from the time the image was presented until the participant terminated the search with a key press. A main effect of viewing condition was found [$F(2, 88) = 69.35$, $p < .001$, partial $\eta^2 = .61$], with Bonferroni-corrected pairwise comparisons indicating significantly longer decision times for image preview relative to free search ($p < .001$) and longer decision times for eye movement preview in comparison to both image preview and free search (both ps < .001). The important comparison across these conditions is between the image preview and eye movement preview conditions since these both required participants to view each image for a minimum of 20s - as distinct from the control condition where this 20s viewing requirement did not arise. No difference in decision time was found for participant expertise [$F(1, 44) = 1.36$, $p > .20$, partial $\eta^2 = .03$] and there was no interaction between participant expertise and viewing condition [$F(2, 88) = .06$, $p > .20$, partial $\eta^2 < .01$]. Model expertise showed neither a main effect nor any interactive effects with other factor [all Fs < 1.48, all ps > .23].

Overall our results indicate that both groups of radiographers demonstrated significantly improved performance when shown another person's eye movements. The improvement by novices was expected because their underdeveloped visual search

behavior has been linked to their relatively poor performance in x-ray image analysis (Kundel et al., 1978). What was not expected was that experienced radiographers would also benefit from being shown another person's eye movements. It was predicted that experienced radiographers would have already encountered enough cases to have developed effective scanning and search strategies, such that knowing where another person looked would only distract them from implementing these strategies. The lack of a performance difference in the eye movement conditions may indicate that: (1) insufficient cases were presented to enable experience related differences to emerge; or (2) that the experienced radiographers had not encountered enough chest x-ray images in their professional practice to develop the aforementioned search strategies. Thus, although experienced radiographers were generally more proficient than novice radiographers, any declining utility of gaze following may have been more apparent with expert radiologists specialized in chest x-ray interpretation.

Another finding relating to observer performance concerned the interaction between participant expertise and viewing condition, whereby novice performance was worse in the image preview condition compared with that of experienced radiographers. This effect may have arisen as a negative consequence of increasing the time available for the novices to examine presented images in the absence of any useful visual cues. This interpretation is consistent with previous studies showing that extended viewing of x-rays often leads to an *increase* in the number of false positives (Christensen et al., 1981; Edwards, Ricketts, Dubbins, Roobottom, & Wells, 2003; Manning et al., 2004), whereas faster decision times are typically an indicator of expertise (Nodine et al., 2002).

In the presence of *useful* visual cues, however, it may well be that longer decision times are associated with performance gains. This is precisely what was found in the eye movement preview condition, where both groups of radiographers took longer to perform

the task but also had the highest figure-of-merit scores. We propose that more than just visual search is involved in the eye movement preview condition in our study. For example, in addition to the primary task of identifying nodules, radiographers in this condition also had to interpret where another person looked in relation to each image and determine whether this information was useful in helping them make their own decisions. In this way, there were multiple cognitive demands associated with interpreting another person's eye movements that would have been likely to require longer decision times.

Intriguingly, Experiment 1 revealed no difference between those radiographers watching an expert search for nodules and those radiographers watching a novice search for nodules. Although Nalanagula et al. (2006) have shown that novices improve when shown an expert's eye movement patterns in an eye movement feed-forward paradigm, the present findings indicate that this improvement can also be obtained from viewing a novice's eye movement patterns. There is an implicit assumption that improved performance will be confined to the eye movement patterns of experts, because an expert should be more likely to lead observers to the most task-relevant areas (e.g., Kundel & Nodine, 1983; Nodine et al., 2002). However, it is not the case that novices totally ignore task-relevant areas. On the contrary, Table 2 shows that the distribution of fixations at nodule locations was similar for both the expert and novice models (collapsed across images). Therefore, the novice's eye movements may have also led observers to consider these areas. Providing that the novice was looking at a nodule, even if the novice did not make the correct decision or was less confident than the expert, the novice's eye movements could have still been useful in guiding observers to consider these areas.

<< Insert Table 2 about here >>

In line with this latter argument it seems that the primary factor in learning from another's gaze relates to the *task-specificity* of the eye movements, rather than the

expertise of the individual who provided them (see Land, 2006; Yarbus, 1967, for

evidence supporting the task-specific nature of eye movements). Indeed, we suggest that

if an observer viewed eye movements that were unrelated to the nodule-detection task

then there should be *no* performance benefit because these eye movements would not

guide attention towards task-relevant areas. This proposal, however, requires empirical

validation, as it is also possible that the performance benefit observed in the eye

movement preview condition arose simply because this condition elevated the general

attention or arousal of participants, thereby leading to more aggressive scanning of

images. Experiment 2 directly addressed these issues by comparing the effects of viewing

an expert's task-specific eye movement patterns versus a set of eye movement patterns

that were unrelated to the nodule-detection task.

A final issue to contend with is how the radiographers perceived the model. In

Experiment 1 the identity and level of expertise of the model were kept ambiguous.

Research has shown that integrating advice from others into the decision-making process

can involve some degree of authority bias and is influenced by the observer's own level

of experience, and the experience of the advisors (Harvey & Fischer, 1997; Schrah, Dalal,

& Sniezek, 2006). Although it is conceivable that these adjustments may occur in the

present study, we would still argue that the *task-specificity* of the eye movement behavior

would remain as the primary factor in learning from another's eye movements. That is,

explicitly knowing that the search behavior came from an expert might well influence the

uptake of potentially *helpful* information, but in the case of *unhelpful* information (e.g., in

the form of unrelated eye movement patterns) such knowledge should be of little value.

Experiment 2

Experiment 2 aimed to establish whether the performance improvement in

Experiment 1 was caused by the presentation of task-specific eye movement patterns or

whether it was a consequence of providing dynamic visual stimuli during image preview that led to a general increase in attention or arousal. To arbitrate between these alternative accounts Experiment 2 involved three conditions: (1) image-preview control, which was not associated with any eye movement cues; (2) "expert-search" preview, which used the expert eye movement patterns from Experiment 1 as visual cues; and (3) "unrelated" preview, which used eye-movement patterns related to a different visual task as visual cues (i.e., an observer's free search of features within fractal images). In line with previous evidence for the task-specific nature of eye-movements (Land, 2006; Yarbus, 1967) we predicted that performance improvements in Experiment 2 would arise only in the expert-search preview condition, even though participants in both eye movement conditions were explicitly told that they would be following an expert radiologist.

Method

Participants. Sixty participants were selected for the study: 30 first-year radiographers (8 male; 22 female) and 30 experienced radiographers (10 male; 20 female) with a minimum of five years experience interpreting medical images from a range of modalities. Participants had normal or corrected-to-normal vision and were each paid £6 for completing the experiment.

Materials. Forty chest x-rays were selected from the test bank described in Experiment 1. Twenty images were normal and contained no nodules and 20 were abnormal and contained up to 5 nodules in each image, with a total of 46 nodules across the 20 abnormal images. Images were standardized to fit an LCD display (resolution of 1280 x 1024) and were presented in a random order using ClearView 2.6.

Design. The experiment employed a 2 x 3 between participants design with two factors: participant expertise, with two levels (novice vs. experienced); and viewing condition, with three levels (image preview; expert-search preview; unrelated preview).

Participants were randomly assigned to one of the three viewing conditions. Image

preview was identical to Experiment 1, requiring participants to view each x-ray for 20s

before being allowed to mark the location of any nodules. The expert-search preview was

the same as the eye movement preview in Experiment 1, with the exception that all the

eye movement recordings were from the highest performing expert radiologist (i.e., no

novice eye movements were included). To obtain the unrelated preview a non-

radiographer was asked to look at 40 fractal images for 20s each. The individual was free

to examine the fractals in any manner whilst eye movements were recorded using a Tobii

x50 eye tracker. The fractal images varied in complexity and spatial configuration, and it

was assumed that the non-radiographer would look at whatever features within each

image attracted attention. The resulting dynamic eye movement pattern for each fractal

was extracted and superimposed over a randomly selected chest x-ray using AVI video

editing software. This produced an identical set of chest x-rays to the expert-search

preview condition, with the only difference being that the eye movements in the unrelated

preview condition had no direct association to the chest x-ray image (i.e., it was purely by

chance whether the gaze cursor landed on a nodule).

Table 3 shows the distribution of fixations for the unrelated and expert-search eye

movements at nodule locations. It is evident that all metrics for the expert-search eye-

movements show higher values at nodule locations than for the unrelated eye movements.

These data attest to the task-relevance of the eye movement patterns. It is also apparent

that at least some fixations fall at nodule locations in the unrelated preview condition.

<< Insert Table 3 about here >>

Procedure. Participants had to identify nodules following the same procedure as

Experiment 1. Participants in the image preview condition viewed each image for 20s

before making a decision. Participants in the unrelated and the expert-search preview

conditions were told they would be "shown an eye movement recording indicating where an expert radiologist examined each image." As such, participants in the unrelated preview condition believed they were following an expert's eye movement patterns. As with Experiment 1, participants in the unrelated and the expert-search preview conditions were not told whether the expert's diagnosis for each image had been correct.

Results and Discussion

The figure-of-merit scores obtained from the JAFROC analysis showed a main effect of viewing condition [$F(2, 54) = 8.53$, $p < .001$, partial $\eta^2 = .24$], with performance significantly better with expert-search preview compared to image preview and unrelated preview (see Table 4). Performance between novice and experienced radiographers was significantly different, with experienced radiographers outperforming novices [$F(1, 54) = 12.00$, $p < .001$, partial $\eta^2 = .18$]. A reliable interaction was also found between participant expertise and viewing condition [$F(2, 54) = 4.07$, $p < .023$, partial $\eta^2 = .13$].

<< Insert Table 4 about here >>

Simple main effects analyses were undertaken for each level of expertise. Novice radiographers showed a reliable performance difference across conditions [$F(2, 27) = 12.53$, $p < .001$, partial $\eta^2 = .48$]. Bonferroni-corrected pairwise comparisons confirmed that the expert-search preview condition led to higher performance than the image preview and unrelated preview conditions (both ps $< .001$). The unrelated preview and image-preview conditions did not differ ($p > .20$). Experienced radiographers did not show a reliable difference in performance across conditions [$F(2, 27) = 0.47$, $p > .20$, partial $\eta^2 = .03$]. These analyses indicate that only novice radiographers benefited from viewing an expert's search related eye movement patterns before making a decision regarding nodules. Simple main effects analyses were undertaken for each level of viewing condition. Experienced radiographers were significantly better than novice

radiographers in the image preview condition [$F(1, 19) = 7.15$, $p < .016$] and the unrelated preview condition [$F(1, 19) = 11.84$, $p < .003$] but not in the expert-search preview condition [$F(1, 19) = 0.13$, $p > .20$]. Overall, these analyses clarify that the novice and experienced radiographers only performed equivalently when eye movement cues were task related. Furthermore, these findings support the claim that viewing another's search behavior aids performance by guiding observers' towards task-relevant areas. The alternative hypothesis – that performance benefits arise simply because eye movement cues elevate participants' general attention or arousal – received no support.

All decision times for the three viewing conditions in Experiment 2 included 20s of preview time (see Table 4). The distribution of decision times indicated that one novice radiographer was over 2.5 SD from the mean; this participant's data were removed from the analysis. There was no main effect of viewing condition [$F(2, 53) = 1.67$, $p > .20$, partial $\eta^2 = .06$], however, there was a main effect of expertise [$F(1, 53) = 4.04$, $p < .049$, partial $\eta^2 = .07$), with novice radiographers taking longer than experienced radiographers. No interaction was found between expertise and viewing condition [$F(2, 53) = .81$, $p > .20$, partial $\eta^2 = .03$]. These results indicate that the experienced radiographers were faster than novices in diagnosing each image by 7.67s on average.

Viewing task-related eye movement patterns once again facilitated performance, but only for novice radiographers. The fact that this improvement was not consistently found by experienced radiographers suggests this benefit may be short-lived and restricted to the early stages of learning. In contrast to Experiment 1, following and interpreting task-specific eye movements was not accompanied with increased decision times relative to the image preview condition. This lack of a difference may reflect the fact that the number of trials per condition increased from 14 in Experiment 1 to 40 in Experiment 2. As a result, participants in Experiment 2 had more time to become familiar

with interpreting the model's eye movements and to learn how to integrate another's search behavior with their own decision-making. The latter methodological difference may also explain why experience-related differences in JAFROC scores were more noticeable, despite the smaller number of participants per condition in Experiment 2.

A key finding from Experiment 2 was that viewing unrelated eye movement patterns did not lead to performance benefits, despite participants being told they would be following where a consultant radiologist looked. Hence the improvement in the expert-search preview condition was not simply due to a general increase in attention or to the observer's knowledge that an expert was helping them. Instead, following task-specific eye movements must have provided some additional information as to how to interpret each x-ray. Understanding what separates the eye movement patterns of the expert search and unrelated conditions should reveal what this additional information might be.

Table 3 shows that the expert-search model looked at nodules more often and for longer, whereas in the unrelated preview condition, gaze position coincided with nodules less frequently and for shorter duration. As such, it is likely that observers in the expert-search preview condition were guided towards nodules more often than observers in the unrelated preview condition and that it is this increase in attentional guidance towards task-relevant areas that explains the performance benefit. Similarly, where the eye movement metrics indicated no difference in attentional guidance, for example, between the expert and novice model in Experiment 1, we found no difference in performance.

Experiment 3

The controversial findings from Experiment 1 implied that the expertise of the model was not a contributing factor in benefitting from another's eye movements, however, it is not clear whether this reflected any outstanding methodological issues. To address these concerns, Experiment 3 employed a between-participants design with four

preview conditions: (1) "naïve-no-task", which used the eye movement patterns of a naïve non-radiographer with no previous experience in medical image perception examining each image under free search without any task; (2) "naïve-search", which used the eye movement patterns of another naïve non-radiographer with no previous experience in medical image perception, but who was asked to identify nodules using just the practice examples as a guide of what to look for; (3) "expert-search", which used the exact same eye movement patterns of the expert model from Experiment 2; and (4) "incongruent-search", which also used the exact same eye movement patterns of the expert model from Experiment 2 but superimposed these dynamic eye movements over a different chest x-ray from the same test bank.

These four conditions allowed us to make a series of predictions. Experiment 2 confirmed that task-specificity of eye movements is a major part of benefiting from viewing another's eye movements, so we can expect that viewing naïve-search and/or expert-search will lead to performance improvements compared to naïve-no-task. Furthermore, if viewing naïve-search and expert-search *both* lead to an improvement in performance over the naïve-no task condition, this would support our earlier findings that the expertise of the model is not a major part of benefiting from viewing another's eye movements, especially as the experience gap between these models is substantial. The incongruent-search condition is unique in that it represents typical expert search behavior, which is task-specific, but not image specific. Therefore, this condition would involve the same periods of extended gaze (a notable factor in communicative gaze, Kleinke, 1986; Velichkovsky, 1995), but attentional guidance towards nodules would be less than in the expert-search preview condition. As a result, it is expected that viewing this model would not lead to a benefit in performance over the naïve-no-task condition.

Method

Participants. Forty first-year radiographers (14 male; 26 female) were selected for the study given that the benefit of following another's gaze primarily occurs during the *early* stages of learning. Participants had normal or corrected-to-normal vision and were each paid £6 for completing the experiment.

Materials. The same 40 chest x-rays were selected from the test bank described in Experiment 2 (i.e., 20 normal images and 20 abnormal images), with a total of 46 nodules across the 20 abnormal images. Images were standardized to fit an LCD display (resolution of 1280 x 1024) and were presented in a random order using ClearView 2.6.

Design. Observer performance was investigated across four between-participant conditions (naïve-no-task; naïve-search; incongruent-search; expert-search). Participants were randomly assigned to one of the four viewing conditions. To obtain the naïve-no-task preview, a naïve non-radiographer was asked to look for 20s at each of the 40 x-rays under free search while eye movements were recorded. Because the naïve observer had no experience of medical image perception or the nodule detection task, it was assumed that this naïve individual's eye movements would be driven simply by exogenous factors, with gaze directed towards whatever features within each image attracted attention. To obtain the naïve-search preview, a second naïve non-radiographer was given five practice examples of what nodules look like and was asked to identify nodules from the 40 x-rays under free search while eye movements were recorded. Although this naïve observer also had no previous experience of the nodule detection task, it was assumed that this naïve individual's eye movements would be driven by both exogenous and endogenous factors. As expected, the resulting figure-of-merit for this naïve observer (.51) was much lower than the novice used in Experiment 1 (.66) and the expert used in Experiments 1-3 (.79).

The expert-search preview in this experiment was identical to the expert-search preview in Experiment 2. However, to form the incongruent-search preview, these dynamic eye movement patterns were extracted and superimposed over different chest x-rays using AVI video editing software. This produced an identical set of x-rays to the expert-search preview, with the only difference being that the eye movements in the incongruent-search preview, although typical of nodule search behavior, were not directly associated with the underlying x-ray image. For example, the dynamics of the eye movement behavior (e.g., extended gaze at specific locations) were identical, but were not congruent with the underlying image. The eye movement metrics of the four models are shown in Table 5.

<< Insert Table 5 about here >>

Procedure. Participants had to identify nodules following the same procedure as in Experiments 1 and 2. In all conditions participants were shown a set of dynamic eye movement patterns for 20s before marking the location of nodules and providing confidence ratings. Consistent with Experiment 2, participants in all conditions were told that they would be "shown an eye movement recording indicating where an expert radiologist examined each image." As with the previous experiments, participants were not informed whether the model's diagnosis for each image had been correct.

Results and Discussion

Ratings of true positives and false positives were subjected to a JAFROC analysis as described above. The resulting figure-of-merit scores were then analyzed using a one-way ANOVA with four levels (naïve-no-task; naïve-search; incongruent-search; expert-search), which revealed a main effect of viewing condition, $F(3, 39) = 8.77$, $p < .001$, partial $\eta^2 = .42$. This main effect of viewing condition was analyzed using Scheffé comparisons. As shown in Table 6, performance following naïve-search was better than

naïve-no-task ($p < .037$) and incongruent-search ($p < .021$), but was not different from expert-search ($p > .20$). Similarly, performance following expert-search was better than naïve-no-task ($p < .006$) and incongruent-search ($p < .003$), and there was no difference between naïve-no-task and incongruent-search ($p > .20$).

<< Insert Table 6 about here >>

It seems that novice radiographers viewing a naïve model searching for nodules led to comparable benefits in performance to when they viewed an expert model searching for nodules. Although viewing the expert model led to slightly higher performance than when viewing the naïve model (M = .75, SD = .05 vs. M = .72, SD = .09) this difference was not significant. Experiment 3 tried to widen the experience gap between the models by using a naïve non-radiographer, rather than a novice radiographer, however, the results still support earlier findings, that model expertise does not contribute to performance benefits. This observation underlines the likely involvement of task-specific eye movements (rather than expertise-driven eye movements) in engendering performance benefits. This proposal is further supported by the finding (as predicted) that incongruent-search did not lead to performance benefits. The incongruent-search served as a better control than the unrelated preview as it involved the same temporal dynamics as the expert-search preview, including the same periods of extended gaze (e.g., arising when a nodule was found). However, in incongruent-search the gaze cursor landed at nodules less often than in the expert-search preview condition, therefore providing fewer useful cues to guide the observers' attention towards these areas. In sum, if benefits are to occur, eye movements must match both the task and the underlying image.

All decision times for the viewing conditions in Experiment 3 included the 20s of preview time (see Table 6). The decision times were subjected to a one-way ANOVA with four factors (naïve-no-task; naïve-search; incongruent-search; expert-search),

however, there was no main effect of viewing condition, $F(3, 39) = 0.25$, $p > .20$, partial

$\eta^2 = .02$. Consequently, Experiment 3 revealed a marked performance benefit for novice

radiographers viewing a naïve model or an expert model, but this was not accompanied

by a change in the time taken to diagnose each chest x-ray image. This is consistent with

the findings of Experiment 2, and suggests that increased exposure to gaze cues in

Experiments 2 and 3 may have mitigated the decision-time effect seen in Experiment 1.

General Discussion

The present study examined whether viewing another person's search behavior

improved visual search and decision-making during a nodule detection task. Viewing

where another person looked during nodule detection improved subsequent identification

of nodules, however, only novice radiographers demonstrated consistent improvements in

performance for both Experiment 1 and Experiment 2. The increase in figure-of-merit

scores indicated that first year radiographers with limited clinical experience could

perform at the same level as much more experienced radiographers simply by observing

where another person looked for nodules. Although this form of attentional guidance was

linked to the observer's level of expertise, no detrimental effects on performance were

observed for either group of radiographers, which would suggest this method did not

appear to restrict any pre-existing skills. Consequently, the present findings are only

partially consistent with previous expertise reversal studies (e.g., Kalyuga, 2007). On the

whole, it seems that experienced radiographers were already capable of finding the task-

relevant areas in each image and making the correct decision. The fact that novices were

able to improve in this regard highlights some of the skills that they still lack and points

towards the different training requirements for the two experience groups.

In line with our predictions, Experiment 2 established that task-specific eye

movements were necessary for performance improvements to occur. When novice

radiographers followed eye movements that were completely unrelated to the task, they failed to perform better than those in the image preview condition. In addition, this need for task-specific eye movements outweighed the need to follow an expert's eye movements. In Experiment 1, both groups of radiographers demonstrated an improvement in performance irrespective of whether they were shown the search behavior of an expert or a novice. When we re-examined this issue in Experiment 3, even viewing a naïve model searching for nodules led to equivalent improvements. To explain this finding, it is first worth mentioning that even the expert model did not identify all the nodules, and therefore, could not have acted as a perfect aid for observers. Nonetheless, the expert model and the naïve model searching for nodules clearly performed differently according to their respective figure-of-merits, yet their respective eye movement metrics showed substantial overlap at nodule locations (see Table 5). As with the novice model in Experiment 1, the naïve model performed worse than the expert, either by making too many errors or by not being as confident as the expert when making correct decisions. Nevertheless the naïve model's search behavior could still have been useful in guiding other novice observers to (re)consider task-relevant areas. Knowing where another person looks influences our own perceptions (Becchio et al., 2008). Unfortunately, the results do not distinguish whether the gaze cues highlighted task-relevant areas that would have otherwise been missed or whether the gaze cues raised the saliency of areas that the observers would have considered anyway. Thus, future studies will need to differentiate how gaze cueing independently affects the accuracy and confidence of subsequent decisions.

What did become clear was that the benefit of viewing another's search behavior was not a result of a speed/accuracy tradeoff. Although viewing another's search behavior led to longer decision times in Experiment 1, when the number of images per

condition was increased in Experiments 2 and 3, we found improvements in performance without an increase in decision times. Given that this was the first time radiographers would have seen another's search behavior in this way, it is perhaps not surprising that with sufficient examples radiographers were soon able to integrate this additional information without prolonging decision times. Indeed, it was only with sufficient images in Experiment 2 that experienced radiographers reached decisions faster than novice radiographers, which is what one would typically expect from the two experience groups (Nodine & Mello-Thoms, 2000). Yet despite these discrepancies, once again, model expertise had no effect on decision times in any of the experiments.

Given that the model's expertise was not a major factor in facilitating performance we need to reconcile this finding with the evidence that an expert's eye movement patterns are more efficient than novices. Until now, the few studies that have shown another's eye movements to observers (e.g., Nalanagula et al., 2006; Sadasivian et al., 2005; Stein & Brennan, 2004; Van Gog et al., 2009; Velichkovsky, 1995) have not examined the interaction between observer expertise and model expertise. As a result, there is little evidence to suggest at what point this form of gaze following ceases to be useful, or whether similar performance gains could arise from following a novice's eye movements. It is encouraging that more studies are examining collaborative visual search, yet the assertion that expert eye movements should be especially useful in training novices where to look neglects the possibility that a *peer's* eye movements could also be useful in directing attention (see Brennan et al., 2008; Stein & Brennan, 2004; see also Hinds, Patterson, & Pfeffer, 2001). That is, even the extensive literature on gaze perception has yet to determine if we follow another's gaze based on whether the person is an "expert" or "novice." Instead, current research indicates that our tendency to use another's gaze relies on factors such as trustworthiness, consistency and prior helpfulness

(Bayliss & Tipper, 2006). It is noteworthy that outside the gaze perception literature we find clear examples of how observers differ in the way they respond to guidance from experts or novices (Harvey & Fischer, 1997; Schrah et al., 2006), or for that matter, computer generated advice (Önkal, Goodwin, Thomson, Gönül, & Pollock, 2009).

Part of the problem we have highlighted is that we are dealing with two issues: (1) whether the actual eye movement patterns of experts and novices are different enough to lead to differential benefits; and (2) how we differentially perceive experts and novices in general. The present study tried to control the latter so that any experience-related differences in eye movement patterns would manifest in different effects on performance. To do this, observers were not told the identity of the model in Experiment 1, or were told that they were following an expert in Experiments 2 and 3. This approach did not lead to differential effects on performance based on model expertise, but did lead to differential effects based on the task-specificity of the eye movement patterns. Therefore, had observers been told that they were following a naïve or novice model rather than an expert model, this may have indeed led to different effects on performance. However, such effects would have been driven by biases in the perception of the model rather than any differences in the eye movements shown.

With these issues accounted for, it may be that this preference to using an expert's search behavior as a decision aid involves an understated scaling issue. Namely, most of the expert/novice differences in eye movement patterns are obtained from statistically analyzing a series of eye movement metrics, which typically deem an expert as looking at task-relevant areas quicker and more often than novices (Chapman & Underwood, 1998; Charness et al., 2001; Krupinski, 1996). However, it is important to consider whether these expert/novice differences in eye movements are, in fact, relatively minor in comparison to the sheer variability in eye movements that arise because of task-specific

demands (Land, 2006; Yarbus, 1967). In other words, the scale of these experience-related differences in eye movements is arguably much smaller than the magnitude of task-related differences in eye movements. Therefore, although these small visual-perceptual differences may undoubtedly play an important role in the expert's enhanced recognition of task-specific features and allow for improved decisions, it appears that these differences are too subtle to be transferred through observation, or more specifically, gaze following. Instead, we found that task-specific attentional guidance via another's search behavior can be useful during the early stages of learning, regardless of model expertise, whereas we found no performance benefit in the instances where these eye movements were not completely congruent with the task and the underlying image. Consequently, new studies continuing this preference for expert eye movements (e.g., Jarodzka, Scheiter, Gerjets, & Van Gog, 2010) should consider carefully the task-specificity of eye movements when shown to other observers.

It should also be made clear that even though task-related eye movement patterns provided useful cues in the present study, it is essential that these same cues do not mislead observers when no nodules are present. For example, half of the images that radiographers examined contained no nodules, but if the model made a prolonged gaze at a particular location but was actually examining a normal image without any nodules, this could have misled observers following this gaze cue. This potentially negative effect would, however, have been reflected in a diminished figure-of-merit score, even in the expert-search preview condition, but as we found an *improved* figure-of-merit score in the search-related conditions, the overall effect of these particular gaze cues was positive.

Moreover, in the instances where the eye movement patterns *did* converge on nodule locations, the eye movement recordings were not manipulated to highlight this fact through color-coding cues (e.g., Nalanagula et al., 2006). Whilst it is encouraging

that we found improvements without having to manipulate gaze cues in this way, this additional manipulation may underlie how the novices in the Nalanagula et al. study were able to transfer what they had learnt to new circuitry boards. Thus, one of the limitations of the present study is that it did not examine whether following another's search behavior could lead to similar transfer effects. Clearly, this issue would need to be resolved if this method were to be developed as part of a larger scale training intervention.

By taking into account the interaction between observer expertise and model expertise, this investigation has revealed new insights into attentional guidance and the role of joint attention during real-world visual search. In particular, our results emphasize the different training requirements of observers, especially as they become more experienced. For example, whereas the present study explored how attentional guidance via another's search behavior can affect observer performance, it appears that the associated benefits may be short-lived and restricted to the early stages of learning. In contrast, with regard to improving expert performance, it is still the developments in computer-aided-detection (CAD) software that have had the most success in helping frontline practitioners minimize the number of erroneous decisions (Doi, 2005; though see Fenton et al., 2007). For this reason, we are not advocating that the paradigm described here would be appropriate for experts in the field of medical image perception. Instead, we point out that the development of effective visual search and decision-making strategies is an important component in the early learning phases of medical image perception. Nevertheless, until novices can develop these skills on their own, they can still make use of another's search behavior to scaffold their nodule detection performance.

Author Note

References

Bard, E. G., Anderson, A. H., Chen, Y., Nicholson, H. B. M., Havard, C., & Dalzel-Job, S. (2007). Let's you do that: Sharing the cognitive burdens of dialogue. *Journal of Memory & Language, 57*, 616-641.

Bayliss, A. P., Frischen, A., Fenske, M. J., & Tipper, S. P. (2007). Affective evaluations of objects are influenced by observed gaze direction and emotional expression. *Cognition, 104*, 644-653.

Bayliss, A. P., & Tipper, S. P. (2006). Predictive gaze cues and personality judgments: Should eye trust you? *Psychological Science, 17*, 514-520.

Beam, C. A., Krupinski, E. A., Kundel, H. L., Sickles, E. A., & Wagner, R. F. (2006). The place of medical image perception in 21st-century health care. *Journal of the American College of Radiology, 3*, 409-412.

Becchio, C., Bertone, C., & Castiello, U. (2008). How the gaze of others influences object processing. *Trends in Cognitive Sciences, 12*, 254-258.

Brennan, S. E., Chen, X., Dickinson, C. A., Neider, M. B., & Zelinsky, G. J. (2008). Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition, 106,* 1465-1477.

Brooks, R., & Meltzoff, A. N (2005). The development of gaze following and its relation to language. *Developmental Science, 8*, 535-543.

Butterworth, G., & Jarret, N. (1991). What minds have in common in space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology, 9,* 55-72.

Chakraborty, D.P. (2000). The FROC, AFROC, and DROC variants of the ROC analysis. In J. Beutel, H. L. Kundel, & R. L. Van Metter (Eds.), *Handbook of Medical Imaging, Vol. 1: Physics and Psychophysics* (pp. 771-796). Bellingham, WA: SPIE.

Chakraborty, D. P. (2006). Analysis of location specific observer performance data: Validated extensions of the jackknife free-response (JAFROC) method. *Academic Radiology, 13*, 1187-1193.

Chakraborty, D. P., & Berbaum, K. S. (2004). Observer studies involving detection and localization: Modeling, analysis and validation. *Medical Physics*, *31*, 2313-2330.

Chapman, P. R., & Underwood, G. (1998). Visual search of driving situations: Danger and experience. *Perception, 27*, 951-964.

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.

Christensen, E. E., Murry, R. C., Holland, K., Reynolds, J., Landay, M. J., & Moore, J. G. (1981). The effect of search time on perception. *Radiology, 138*, 361-365.

Corkum, V., & Moore, C. (1995). Development of joint visual attention in infants. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 61-83). Hove, England: Erlbaum.

Doi, K. (2005). Current status and future potential of computer-aided diagnosis in medical imaging. *The British Journal of Radiology, 78*, S3–S19.

Donovan, T., & Manning, D. J. (2006). Successful reporting by non-medical practitioners such as radiographers, will always be task-specific and limited in scope. *Radiography, 12,* 7-12.

Donovan, T., Manning, D. J., & Crawford, T. (2008). Performance changes in lung nodule detection following perceptual feedback of eye movements. In B. Sahiner, & D. J. Manning (Eds.), *Proceedings of SPIE Medical Imaging 2008: Image Perception, Observer Performance, and Technology Assessment, 9* (pp. 6917-03/1-9). San Diego, CA: SPIE.

Donovan, T., Manning, D. J., Philips, P. W., Higham, S., & Crawford, T. J. (2005). The

effect of feedback on performance in a fracture detection task. *Progress in*

Biomedical Optics and Imaging (SPIE), 6, 79-85.

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999).

Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition, 6*, 509-

540.

Duncker, K. (1945). On problem solving. *Psychological Monographs, 58*(5, No. 270).

Edwards, A. J., Ricketts, C., Dubbins, P. A., Roobottom, C. A., & Wells, I. P. (2003).

The effect of reporting speed on plain film reporting errors. *Clinical Radiology*, *58*,

971-979.

Fenton, J. J., Taplin, S. H., Carney, P. A., Abraham, L., Sickles, E. A., D'Orsi, C., et al.

(2007). Influence of computer-aided detection on performance of screening

mammography. *The New England Journal of Medicine, 356*, 1399-1409.

Fitzgerald, R. (2001). Error in radiology. *Clinical Radiology, 56,* 938-946.

Flom, R., & Pick, A. D. (2007). Increasing specificity and the development of joint visual

attention. In R. Flom, K. Lee, & D. Muir (Eds.), *Gaze-following: Its development and*

significance (pp. 95-111). London: Erlbaum.

Gibson, J., & Pick, A. (1963). Perception of another person's looking behavior. *American*

Journal of Psychology, 76, 386-394.

Grant, E. R., & Spivey, M. J. (2003). Eye movements and problem solving: Guiding

attention guides thought. *Psychological Science*, *14*, 462-466.

Hanna, J. E., & Brennan, S. E. (2007). Speakers' eye gaze disambiguates referring

expressions early during face-to-face conversation. *Journal of Memory & Language,*

57, 596-615.

Harvey, N., & Fischer, I. (1997). Taking advice: Accepting help, improving judgment, and sharing responsibility. *Organizational Behavior & Human Decision Processes, 7*, 117–134.

Hinds, P. J., Patterson, M., & Pfeffer, J. (2001). Bothered by abstraction: The effect of expertise on knowledge transfer and subsequent novice performance. *Journal of Applied Psychology, 86*, 1232-1243.

Jarodzka, H., Scheiter, K., Gerjets, P., & van Gog, T. (2010). In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learning & Instruction, 20,* 146-154.

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review, 19*, 509-539.

Kleinke, C. L. (1986). Gaze and eye contact: A research review. *Psychological Bulletin, 100*, 78-100.

Krupinski, E. A. (1996). Visual scanning patterns of radiologists searching mammograms. *Academic Radiology, 3*, 137-144.

Krupinski, E. A., Nodine, C. F., & Kundel, H. L. (1993). Perceptual enhancement of tumor targets in chest x-ray images. *Perception & Psychophysics*, *53*, 519-526.

Krupinski, E. A., Tillack, A. A., Richter, L., Henderson, J. T., Bhattacharyya, A. K., Scott, K. M., et al. (2006). Eye-movement study and human performance using telepathology virtual slides: Implications for medical education and differences with experience. *Human Pathology, 37*, 1543-1556.

Kundel, H. L., & Nodine, C. F. (1983). A visual concept shapes image perception. *Radiology, 146*, 363-368.

Kundel, H. L., Nodine, C. F., & Carmody, D. (1978). Visual scanning, pattern recognition and decision-making in pulmonary nodule detection. *Investigative Radiology*, *13*, 175-180.

Kundel, H. L., Nodine, C. F., & Krupinski, E. A. (1990). Computer-displayed eye position as a visual aid to pulmonary nodule interpretation. *Investigative Radiology*, *25*, 890-898.

Land, M. F. (2006). Eye movements and the control of actions in everyday life. *Progress in Retinal and Eye Research, 25*, 296-324.

Litchfield, D., Ball, L. J., Donovan, T., Manning, D. J., & Crawford, T. (2008). Learning from others: Effects of viewing another person's eye movements while searching for chest nodules. In B. Sahiner, & D. J. Manning (Eds.), *Proceedings of SPIE Medical Imaging 2008: Image Perception, Observer Performance, and Technology Assessment, 9* (pp. 6917-15/1-9). San Diego, CA: SPIE.

Lobmaier, J. S., Fischer, M. H., & Schwaninger, A. (2006). Objects capture perceived gaze direction. *Experimental Psychology, 53*, 117-122.

Lobmaier, J. S., Tiddeman, B. P., & Perret, D. I. (2008). Emotional expression modulates perceived gaze direction. *Emotion, 8*, 573-577.

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.

Metz, C. E. (2008). ROC analysis in medical imaging: a tutorial review of the literature *Radiology Physics Technology, 1*, 2-12.

McCarley, J. S., Kramer, A. F., Wickens, C. D., Vidoni, E. D., & Boot, W. R. (2004). Visual Skills in Airport-Security Screening. *Psychological Science, 15*, 302-306.

Nalanagula, D., Greenstein, J. S., & Gramopadhye, A. K. (2006). Evaluation of the effect of feedforward training displays of search strategy on visual search performance. *International Journal of Industrial Ergonomics*, *36*, 289-300.

Nodine, C. F., & Mello-Thoms, C. (2000). The nature of expertise in radiology. In J. Beutel, H. L. Kundel, & R. L. Van Metter (Eds.), *The handbook of medical imaging: Physics and psychophysics* (pp. 859-894). Bellingham, WA: SPIE 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.

Önkal, D., Goodwin, P., Thomson, M., Gönül, S., & Pollock, A. (2009). The relative influence of advice from human experts and statistical methods on forecast adjustments. *Journal of Behavioral Decision Making, 22*, 390-409

Pomplun, M., Ritter, H., & Velichkovsky, B. (1996). Disambiguating complex visual information: Towards communication of personal views of a scene. *Perception*, *25*, 931-948.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372-422.

Sadasivian, D., Greenstein, J. S., Gramopadhye, A. K., & Duchowski, A. T. (2005). Use of eye movements as feedforward training for a synthetic aircraft inspection task. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2005* (pp. 289-300). New York, NY: ACM Press

Schrah, G. E., Dalal, R. S., & Sniezek, J. A. (2006). No decision-maker is an island: Integrating expert advice with information acquisition. *Journal of Behavioral Decision Making, 19*, 43-60.

Stein, R., & Brennan, S. E. (2004). Another person's eye gaze as a cue in solving programming problems. *Proceedings of ICMI 2004, Sixth International Conference on Multimodal Interfaces* (pp. 9-15). New York, NY: ACM Press.

Symons, L. A., Lee, K., Cedrone, C. C., & Nishimura, M. (2004). What are you looking at? Acuity for triadic eye gaze. *The Journal of General Psychology, 131*, 451-469.

Taylor, P. M. (2007). A review of research into the development of radiologic expertise: Implications for computer-based training. *Academic Radiology, 14*, 1252-1263.

Thomas, L. E., & Lleras, A. (2007). Moving eyes and moving thought: On the spatial compatibility between eye movements and cognition. *Psychonomic Bulletin & Review, 14*, 663-668.

Van Gog, T., Jarodzka, H., Scheiter, K., Gerjets, P., & Paas, F. (2009). Attention guidance during example study via the model's eye movements. *Computers in Human Behavior, 25*, 785-791.

Vecera, S. P., & Rizzo, M. (2006). Eye gaze does not produce reflexive shifts of attention: Evidence from frontal-lobe damage. *Neuropsychologia, 44*, 150-159.

Velichkovsky, B. M. (1995). Communicating attention: Gaze position transfer in cooperative problem solving. *Pragmatics and Cognition*, *3*, 199-222.

Vivianni, P. (1990). Eye movements in visual search: Cognitive, perceptual and motor control aspects. In E. Kowler (Ed.), *Eye movements and their role in visual and cognitive processes* (pp. 353-393). Amsterdam: Elsevier.

Yaniv, I., & Kleinberger, E. (2000). Advice taking in decision making: Egocentric discounting and reputation formation. *Organizational Behavior and Human Decision Processes, 83*, 260–281.

Yarbus, A. (1967). *Eye movements and vision*. New York: Plenum Press.

Figure 1.

a) Chest x-ray with three pulmonary nodules highlighted, b) A novice radiographer's current eye movements (upper circle) whilst following an expert radiologist's eye movement patterns (lower circle). Lines from circles indicate previous fixation locations within the last 500ms.
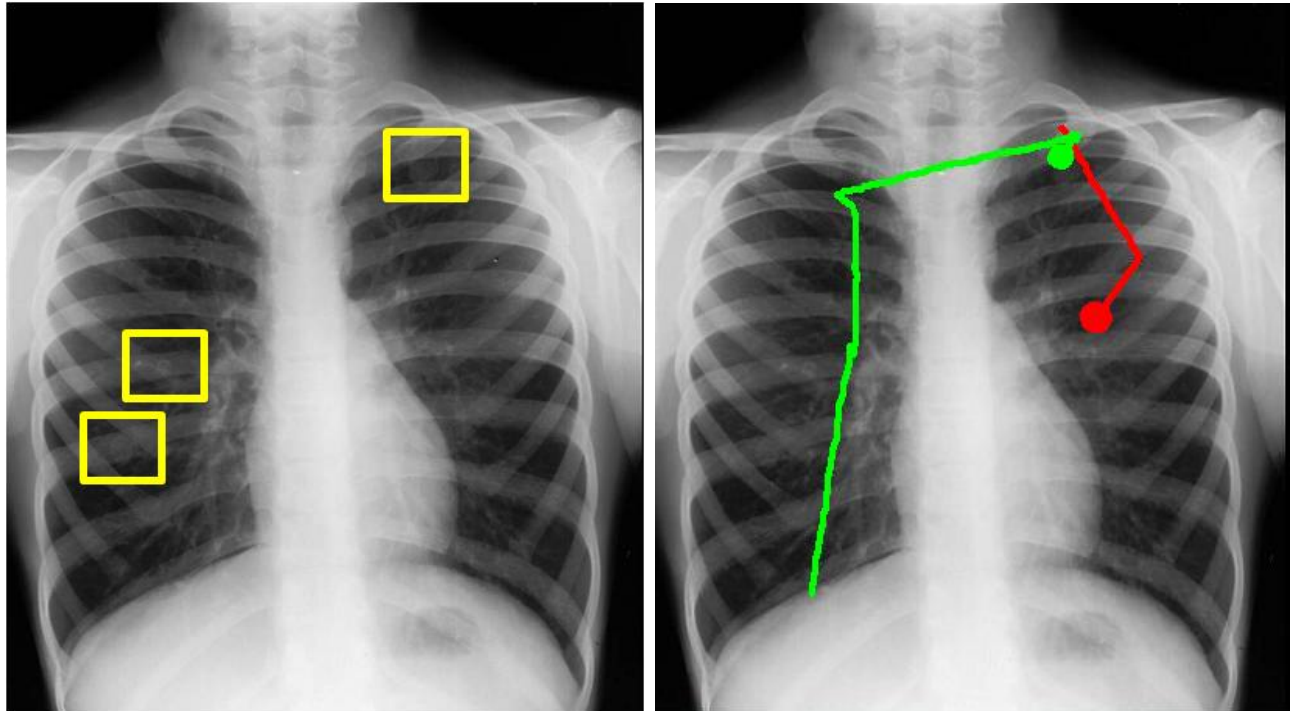
a)   b)

Table 1.

Mean (and Standard Deviation) Performance Measures for Novice and Experienced Radiographers in Experiments 1. Note that for eye-movement preview performance is collapsed over expertise of the model

Observer Expertise	Viewing Condition	JAFROC scores		Decision Time (sec)	
		M	*SD*	*M*	*SD*
Novice	Free search	.56	.04	36.23	9.21
	Image preview	.54	.04	46.25	12.83
	Eye movement preview	.60	.04	56.15	13.07
Experienced	Free search	.57	.04	40.86	17.99
	Image preview	.57	.05	50.88	10.65
	Eye movement preview	.59	.04	59.76	18.63

Table 2.

Mean (and Standard Deviation) Eye Movement Metrics at Nodule Locations for Models

in Experiment 1

Source of Eye Movement Preview	Cumulative Gaze Duration (sec)		Percentage of Gaze Time at Nodules		Average Gaze Duration (msec)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Novice	6.67	5.01	34.30	26.10	772.78	596.00
Expert	7.74	3.24	43.40	17.00	859.79	426.62

Table 3.

Mean (and Standard Deviation) Eye Movement Metrics at Nodule Locations for Models

in Experiment 2

Source of Eye Movement Preview	Cumulative Gaze Duration (sec)		Percentage of Gaze Time at Nodules		Average Gaze Duration (msec)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Unrelated	1.26	1.16	7.10	6.60	305.21	237.37
Expert	5.15	3.27	27.70	17.90	717.15	321.99

Table 4.

Mean (and Standard Deviation) Performance Measures for Novice and Experienced

Radiographers in Experiments 2

Observer Expertise	Viewing Condition	JAFROC scores		Decision Time (sec)	
		M	*SD*	*M*	*SD*
Novice	Image preview	.66	.09	55.35	17.72
	Unrelated preview	.66	.06	60.56	17.02
	Expert-search preview	.79	.04	50.97	10.58
Experienced	Image preview	.75	.06	41.86	14.00
	Unrelated preview	.76	.07	52.64	15.07
	Expert-search preview	.78	.07	49.39	11.06

Table 5.

Mean (and Standard Deviation) Eye Movement Metrics at Nodule Locations for Models

in Experiment 3

Source of Eye Movement Preview	Cumulative Gaze Duration (sec)		Percentage of Gaze Time at Nodules		Average Gaze Duration (msec)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Naïve-no-task	0.74	0.89	3.80	4.50	317.53	325.35
Naïve-search	5.42	2.84	27.80	14.70	1282.45	707.86
Incongruent-search	1.09	1.40	6.10	7.60	230.84	247.08
Expert-search	5.15	3.27	27.70	17.90	717.15	321.99

Table 6.

Mean (and Standard Deviation) Performance Measures for Novice Radiographers in Experiments 3

Viewing Condition	JAFROC scores		Decision Time (sec)	
	M	*SD*	*M*	*SD*
Naïve-no-task	.62	.04	55.03	10.29
Naïve-search	.72	.09	54.14	11.40
Incongruent-search	.61	.10	51.15	11.54
Expert-search	.75	.05	53.04	8.73