

Learning from others: effects of viewing another person's eye movements while searching for chest nodules

Damien Litchfield^{1 a}, Linden J. Ball^a, Tim Donovan^b, David J. Manning^b, Trevor Crawford^a

^a Department of Psychology, Lancaster University, Lancaster LA1 4YF, UK

^b School of Medical Imaging Sciences, Cumbria University, Lancaster LA1 3JD, UK

Abstract

We report a study that investigated whether experienced and inexperienced radiographers benefit from knowing where another person looked during pulmonary nodule detection. Twenty-four undergraduate radiographers (1 year of experience) and 24 postgraduate radiographers (5+ years of experience) searched 42 chest x-rays for nodules and rated how confident they were in their decisions. Eye movements were also recorded. Performance was compared across three within-participant conditions: (1) free search—where radiographers could identify nodules as normal; (2) image preview—where radiographers were first shown each chest x-ray for 20 seconds before they could then proceed to mark the location of any nodules; and (3) eye movement preview—which was identical to image preview except that the 20 second viewing period displayed an overlay of the real-time eye movements of another radiographer's scanpath for that image. For this preview condition half of each group were shown where a novice radiographer looked, and the other half were shown where an experienced radiologist looked. This was not made known to the participants until after the experiment. Performance was assessed using JAFROC analysis. Both groups of radiographers performed better in the eye movement preview condition compared with the image preview or free search conditions, with inexperienced radiographers improving the most. We discuss our findings in terms of the task-specific information interpreted from eye movement previews, task difficulty across images, and whether it matters if radiographers are previewing the eye movements of an expert or a novice.

1. Introduction

As is the case with most domains of expertise, studying the eye movement patterns of groups with different experience levels has provided an insight into how visual search strategies change as a function of expertise. For example, it has been repeatedly shown that experts make more fixations of shorter duration than novices, suggesting that with experience, less time is required to process task-specific information [1,2]. It has also been shown that eye movement metrics which correlate with specific perceptual interpretations (e.g., accurate identification of problem features) can then be used to provide visual cues to direct other observers toward appropriate decision making [3,4]. What is still unclear, however, is whether the actual viewing of eye movement patterns made by an observer can have a direct role in changing their perception. To this end, a handful of studies have recorded an observer's eye movements as they complete a task and subsequently shown the resulting eye movement patterns back to the same observer. In medical imaging, this type of perceptual feedback appears to lead to an improvement in pulmonary nodule detection [5-9] but curiously appears to have a negative impact on fracture detection [10], perhaps due to a lack of task difficulty. In other visual search domains, studies have tended to focus on replaying eye movement patterns of one observer to subsequent observers. For example, using this paradigm in aircraft inspection, it has been shown that detection of faults increases, but so do completion times [11], whereas in program debugging and puzzle problem solving it has been demonstrated that task completion times decrease [12,13]. These mixed effects may be attributed to the different ways eye movement patterns can be presented [14].

The present study expands on this body of research by investigating whether viewing another's eye movements improves visual search and decision-making during a nodule detection task performed by radiographers. Previous

¹ email: d.litchfield@lancaster.ac.uk

work on false negative errors suggests that inadequate visual search can account for up to 30% of errors [15,16]. Novice radiographers often make these scanning errors, but with experience, decision-making and interpretation skills become the source of most errors [17]. Decision-making and interpretation skills require appropriate training and extensive exposure to case studies which is difficult to develop rapidly. In contrast, viewing eye movement patterns derived from another observer completing the same task may help develop scanning and recognition skills, as this would provide novices with an example of an effective search. Therefore it is expected that showing eye movement patterns to novice radiographers will lead to an improvement in decision-making performance. On the other hand, showing eye movement patterns to experienced radiographers may actually have a detrimental effect on diagnostic performance [10], a possibility that links to the “expertise reversal effect” [18], whereby many instructional techniques appear to inhibit the application of pre-existing skills. Based on these assumptions, it is expected that experienced radiographers would show no decision-making improvement or even negative performance when viewing eye movement patterns as this would conflict with existing search strategies. The study will also investigate whether performance is influenced if eye movement recordings shown to other radiographers are derived from a novice radiographer or from an experienced radiologist. This is to test the assumption that viewing an expert’s eye movement patterns leads to greater improvement compared to viewing a less skilled observer.

Our planned focus on decision accuracy as a dependent variable rather than decision time overcomes difficulties arising from the fact that showing pre-task eye movements will inherently increase an observer’s exposure to the image, which will almost invariably promote faster decision times. Since increasing reporting speed generally has no significant effect upon accuracy but can lead to fewer false positives [19], there may be an increase in the number of false positives with increased exposure. To examine such exposure effects head on we have devised three conditions:

1. *Free Search*—Where observers could make decisions regarding nodules immediately.
2. *Image Preview*—Where observers were required to wait for 20s before making a decision.
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 eye movements and half shown an expert’s eye movements.

2. Method and Materials

2.1 Film test bank

From a test bank of 120 anteroposterior chest x-rays, seven normal and seven abnormal images were selected for each of the 3 conditions. Each image set contained the same number of nodules of approximately the same size, shape and conspicuity and were each verified by a consultant radiologist before the study. The images were standardized to fit to an LCD display with a resolution of 1280x1024 and were presented using ClearView 2.6.

2.2 Obtaining the eye movement preview

Five radiologists (experts) and five first-year radiographers (novices) were paid £5 to complete the nodule detection task using this film test bank under a free search condition while their eye movements were recorded using a Tobii x50 eye tracker (Tobii Technology, Stockholm, Sweden). The Tobii x50 is 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 and a fixation radius of 50 pixels. Ratings of true positives and false positives were obtained and subjected to a jackknife free-response receiver operating characteristic (JAFROC) analysis, as this has been shown to yield greater statistical power in discriminating between modalities [20]. The resulting ‘figure of merit’ (FOM) represents the likelihood that a true positive will be given a higher rating than a false positive. The highest and lowest FOMs obtained from the JAFROC analysis were taken as the key criteria for selecting which particular eye movement recordings to show the participants in the main study. The eye movement patterns from the novice with the lowest JAFROC score (0.66) and the radiologist with the highest JAFROC score (0.79) form the stimulus set for the eye movement preview condition in the main study.

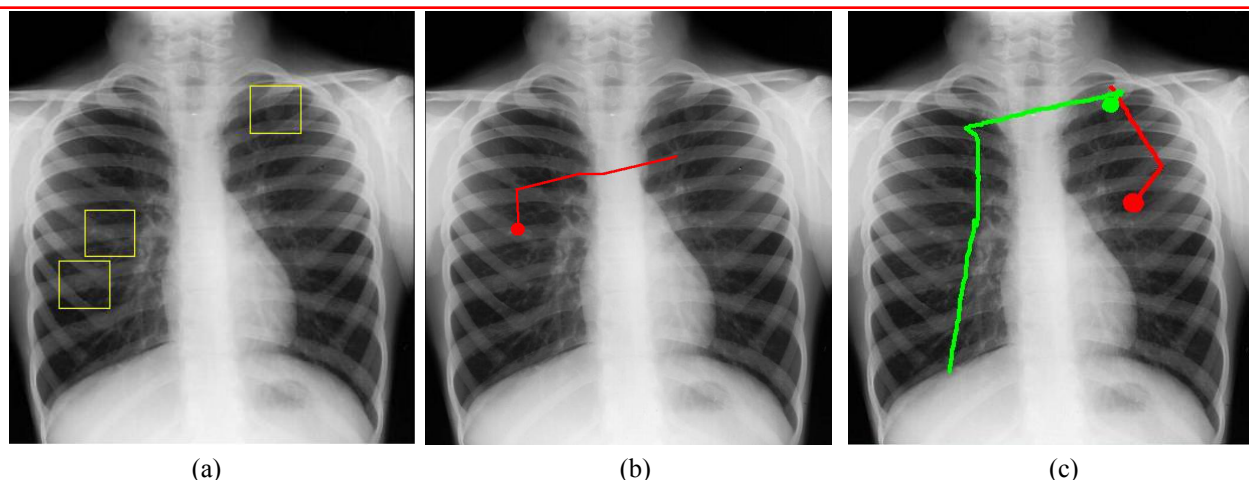


Fig. 1. a) Chest x-ray with three pulmonary nodules highlighted. b) 500ms of pre-recorded eye movement taken from an experienced radiologist as they searched for nodules. c) A novice radiographer's current eye movements (upper circle) as they follow the radiologist's eye movement patterns (lower circle). Lines from circles indicate previous fixation locations.

2.3 Observer groups in the main study

A total of 48 participants, 24 postgraduate radiographers with a minimum of 5 years experience and 24 undergraduate first-year radiography students were selected for the main study. All participants had normal or corrected-to-normal vision and were each paid £6 for completing the experiment.

2.4 Viewing conditions

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. Eye movement preview was identical to image preview except that during the 20s viewing period participants were also shown an overlay of the real-time eye movements of the pre-selected scanpaths for that image (see Fig. 1; for further details see Appendix item 1.). Half of each group was shown the expert's and half were shown the novice's eye movements. In the eye movement preview condition, participants were told that they would be "shown an eye movement recording of someone else examining an image" and that after the recording finished, they would be "presented with the same image and asked to search and identify nodules as before". The only description of the eye movement recordings given to participants was that "the circle represents where someone was looking and the bigger the circle, the longer they were looking at that particular area". The prevalence of nodules across images (50%) was not made known to the participants. No information was given regarding the expertise of the person who had produced the eye movement recording, nor whether the previous person had been correct in their diagnosis.

2.5 Procedure

Participants were seated approximately 60cm from the display. Eye movements were calibrated using a standardised 9-point calibration. Participants were then asked to search for pulmonary nodules from the test bank described previously. Five practice x-ray images were shown before the experiment, demonstrating how nodules differed in size, location, conspicuity, and calcification. When participants located a nodule they used the mouse to indicate the nodule location and clicked the left mouse button to confirm their decision. Once the decision was made, participants were required say how confident they were by calling out a number between 1 and 4, with 4 being "highly confident". Image order was randomised and condition order was completely counterbalanced. The whole experiment took approximately 1 hour to complete.

3. Results

3.1 Data analysis

The experiment employed a 2 x 3 mixed design with two factors: Expertise with two levels (undergraduates with 1 year experience vs. postgraduates with 5 years experience) and Preview with three levels (Free Search; Image Preview; Eye Movement Preview). Ratings of true positives and false positives were subjected to a JAFROC analysis [20]. These performance data and participants' eye movement data were examined using a number of repeated measures ANOVAs. Separate analyses were performed on the type of eye movement preview shown (novice vs. expert).

3.2 Observer performance

The FOMs obtained from the JAFROC analysis showed a main effect of Preview [$F(2, 92) = 30.29, p < .001, \eta^2 = .39$], with performance significantly better with Eye Movement Preview compared to Image Preview and Free Search (see Fig. 2a). This indicates that both undergraduate and postgraduate radiographers were more likely to associate nodules with a higher rating (i.e., make the correct decision) when shown another person's eye movements. However, the magnitude of the difference was relatively small, with only a 3-4 point increase in performance. No difference in performance was found between undergraduate and postgraduate radiographers [$F(1, 46) = 1.95, p = .17, \eta^2 = .04$], however a reliable interaction was found between level of expertise and viewing condition [$F(2, 92) = 4.26, p = .017, \eta^2 = .09$]. Simple main effects analyses were undertaken for each level of expertise. Undergraduate radiographers showed a reliable performance difference across conditions ($F(2, 46) = 25.24, p < .001, \eta^2 = .52$). Bonferroni comparisons confirmed that the Eye Movement Preview condition led to higher scores than the Free Search and Image Preview conditions (both $ps < .001$). Postgraduate radiographers also showed a reliable performance difference across conditions, but with a reduced effect size compared with the undergraduate radiographers [$F(2, 46) = 6.84, p = .002, \eta^2 = .23$]. Bonferroni comparisons confirmed that the Eye Movement Preview condition led to higher scores than the Free Search condition ($p = .034$) and the Image Preview condition ($p = .005$). Simple main effects comparing across expertise levels at each condition revealed no difference for Free Search [$F(2, 46) = 2.13, p = .15, \eta^2 = .04$] or Eye Movement Preview [$F(2, 46) = 0.25, p = .62, \eta^2 = .01$], but a reliable difference for Image Preview [$F(2, 46) = 5.15, p = .028, \eta^2 = .10$]. These analyses clarify that both groups benefit from receiving eye movement patterns before making decision regarding nodules, with the undergraduates showing the greatest performance gains.

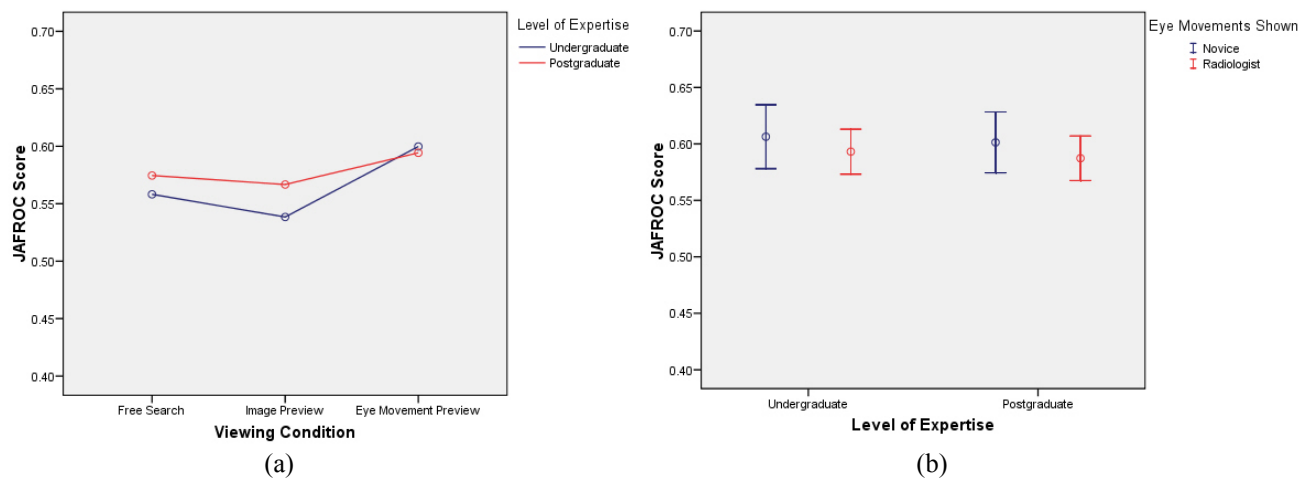


Fig. 2. a) The effects of viewing condition on JAFROC scores for each observer group, and b) performance in eye movement preview condition depending on type of eye movement shown.

3.3 Type of preview cue

Performance in the Eye Movement Preview condition was analysed further by comparing those radiographers shown an expert's eye movements and those shown a novice's eye movements (see Fig. 2b). Results from the analysis of variance revealed no significant difference in performance for either group of radiographers [$F(1, 44) = 1.57, p = .22, \eta^2 = .03$], suggesting that the level of expertise from the eye movement recordings was not a factor in the present study.

3.4 Decision times

The overall decision times for each condition were analysed to examine if Eye Movement Preview increased decision times, as has been seen in previous studies using a similar preview manipulation but in the domain of aircraft inspection [11]. A main effect of Preview was found [$F(2, 92) = 69.31, p < .001, \eta^2 = .60$], with Bonferroni comparisons indicating significantly longer decision times for Eye Movement Preview in comparison to both Image Preview and Free Search (both $ps < .001$) (see Fig. 3.). No difference in decision time was found between undergraduate and postgraduate radiographers [$F(1, 46) = 1.41, p = .24, \eta^2 = .03$] and there was no interaction found between level of expertise and viewing condition [$F(2, 92) = .06, p = .94, \eta^2 < .01$]. The important comparison across these conditions is between Image Preview and Eye Movement Preview as both of these conditions required participants to view each image for a minimum of 20s. What the data clearly indicate is that even though both groups of participants had already viewed the images for 20s in these conditions, participants still took longer (mean = 9.39s) to arrive at a final decision for the eye movement preview condition relative to the image preview condition.

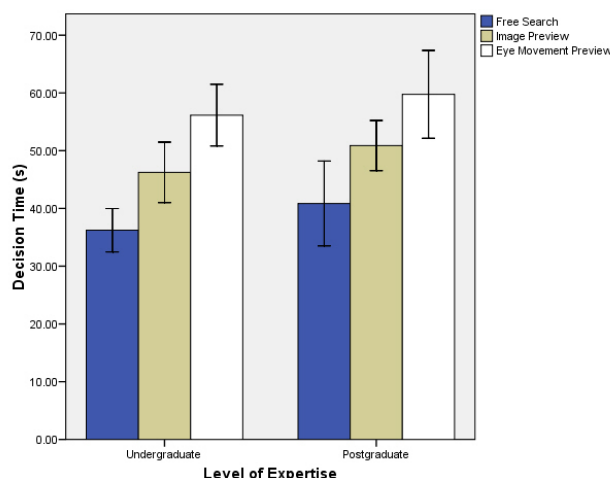


Fig. 3. Average time to diagnose each x-ray image for the three viewing conditions. Note that Image Preview and Eye Movement Preview include an additional 20s of preview time.

3.5 Eye tracking data

Eye-tracking data were gathered for the three conditions to examine the general fixation patterns across viewing conditions. Eye movement data were screened so that only recordings with <25% data loss were analysed, resulting in valid eye tracking data from 18 undergraduate radiographers and 18 postgraduate radiographers. To measure fixation patterns at nodule locations, areas of interest (AOI) were constructed around each nodule with a 5-degree radius from the centre of each nodule. Fixations landing within this area were classified as looking at the nodule. The key eye tracking metrics at nodule locations were:

- 1) Average fixation duration
- 2) Average number of fixations
- 3) Average cumulative dwell time
- 4) Percentage of time spent looking at nodules (total fixation duration at nodules / total viewing time)

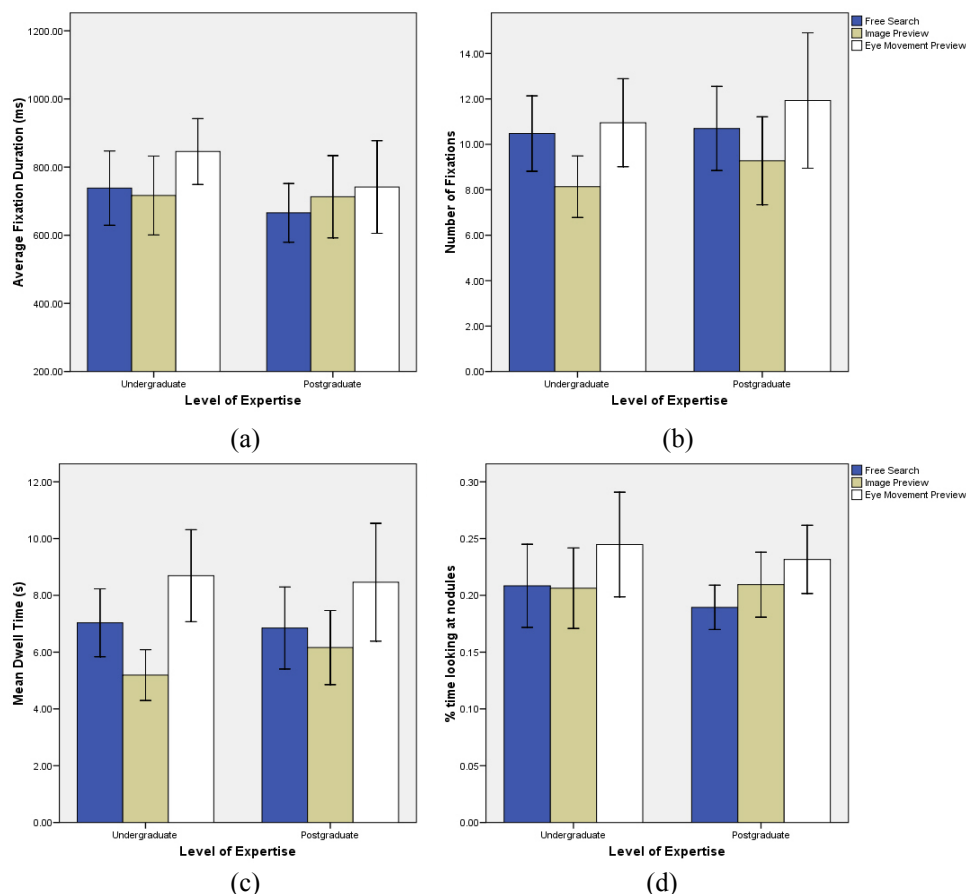


Fig. 4. The fixation patterns at pulmonary nodules of undergraduate and postgraduate radiographers across the three viewing conditions. a) Average fixation duration at nodules, b) number of fixations at nodules, c) cumulative dwell time at nodules and, d) percentage of time spent looking at nodules compared to the total viewing time.

Average fixation duration

A main effect of Preview was found [$F(2, 68) = 5.33, p = .007, \eta^2 = .14$]. Average fixation duration at nodules was longer during Eye Movement Preview than Free Search (see Fig. 4a). No difference in average fixation duration was found between undergraduate and postgraduate radiographers [$F(1, 34) = 0.72, p = .40, \eta^2 = .02$], indicating that the level of processing at nodule locations was not different across expertise groups.

Average number of fixations

A main effect of Preview was found [$F(2, 68) = 7.26, p = .001, \eta^2 = .18$] with significantly fewer fixations made during Image Preview (see Fig. 4b). No difference was found between undergraduate and postgraduate radiographers in the number of fixations made [$F(1, 34) = 0.46, p = .50, \eta^2 = .01$]. The lower number of fixations in Image Preview may relate to the shorter decision times.

Average cumulative dwell time

A main effect of Preview was found [$F(2, 68) = 15.32, p < .001, \eta^2 = .31$], with dwell time significantly different across all conditions (see Fig. 4c). Dwell time was shortest for Image Preview and longest at Eye movement Preview. No difference in dwell time was found between undergraduate and postgraduate radiographers [$F(1, 34) = .048, p = .83, \eta^2 < .01$].

Percentage of time spent looking at nodules

A main effect of Preview was found [$F(2, 68) = 8.94, p < .001, \eta^2 = .20$]. No difference was found between undergraduate and postgraduate radiographers in the percentage of time spent looking at nodules [$F(1, 34) = 0.46, p = .50, \eta^2 = .01$]. Both groups of radiographers spent proportionally more time looking at nodules during Eye Movement Preview than either Free Search or Image Preview (see Fig. 4d).

4. Discussion

This study examined whether showing radiographers another person's eye movements during nodule detection would improve their visual search and subsequent decision-making ability. The results indicate a reliable improvement in performance for both undergraduate and postgraduate radiographers when shown another person's eye movements, with undergraduates showing the largest improvement. This improvement by undergraduates was expected, with the rationale being that viewing another person's eye movements would provide an example of an effective search and therefore help reduce scanning related errors. What was not expected was that postgraduate radiographers would also improve when shown another person's eye movements. It was assumed that postgraduate radiographers would have encountered enough cases to develop their scanning and search strategies so that knowing where another person looked would only distract them from implementing these search strategies. This would have then corresponded to the recurring findings from the expertise reversal effect, which demonstrate that certain learning interventions, such as the one applied in the present study, are only beneficial to those with limited experience [18]. Since this was not the case, it could be that the level of experience between undergraduate and postgraduate radiographers in the present study was not sufficient, which would also explain the lack of experience-related differences from the eye tracking data. One implication of all of this is that negative performance might have been found if radiologists (experts) were tested in the Eye Movement Preview condition.

A closer examination of the decision time data indicates that the improvement for Eye Movement Preview involved both groups of radiographers taking more time to come to a final decision. This is quite remarkable considering that previous research suggests that simply increasing the time spent examining an image often fails to produce an improvement in diagnostic performance [19]. In fact, extended viewing of x-rays often leads to an increase in the number of false positives, which was not the case for Eye Movement Preview in the present study. To explain this contradiction, we suggest that more than just visual search is involved during Eye Movement Preview. For example, in addition to the primary task of identifying nodules in this condition, radiographers also had to interpret where another person looked on each image and determine whether this information was useful in helping them make their own decisions. In this way, there are multiple cognitive demands associated with interpreting another person's eye movements that would be likely to require longer decision times.

There are also at least two factors that could complicate radiographers' interpretation of eye movement patterns during eye movement preview. First of all, it is likely that this would be the first time that the radiographers would have encountered stimuli such as eye movement data and without the benefit of any training in interpreting such information. Only a brief description was given as to what the eye movement recordings represented. Radiographers were also not made aware of key empirical findings, such as the link between average fixation duration and the level of processing. Second, real-time eye movement patterns can be considered as an instrumental cue in that they represent the entire search process, which means that they may also include information relating to unsuccessful search outcomes [12]. In this respect, even if task specific information could be gained from viewing real-time eye movement patterns of other people, these patterns would inherently contain search behaviour that did not indicate a nodule location. In the instances where the eye movement patterns *did* converge on nodule locations, the eye movement recordings were not manipulated to highlight this fact (e.g., through colour-coding cues) as in other studies (e.g., [14]). Yet despite these ongoing issues, the present study still demonstrates that knowing where other people look improves performance on a nodule detection task.

What is more, the results from the present study show that this improvement is maintained whether radiographers are shown the eye movement patterns of an expert or a novice. While studies by Nalanagula, et al., [14] have shown that visual search and performance improve when shown expert eye movement patterns as part of feed-forward training, the present study indicates that improvements may be equally obtained when showing *novice* eye movement patterns

in a feed-forward paradigm. This suggests that the experience-related differences in eye movement patterns replicated in numerous domains may be too subtle to make an observable difference when used as a visual cue. However, an alternative explanation could be that the performance of the novice selected was too high. This could be tested in future studies by selecting recordings based on performance on individual trials rather than showing the eye movement patterns of the novice with the lowest overall performance.

In terms of the eye tracking data, our analyses indicated that more time was spent looking at nodules for the Eye Movement Preview condition than for the Image Preview and Free Search conditions. However, this time-based difference may simply reflect the shorter decision times for the latter conditions, rather than any specific cognitive differences. The measure that we report that captured the percentage of time spent looking at nodules takes this discrepancy into account, and indicates that there is still a increase in the proportion of time spent looking at nodules for the Eye Movement Preview condition. Overall, then, there seems to be supportive evidence for the view that pulmonary nodules received more attention in the Eye Movement Preview condition than Image Preview or Free Search conditions. It is possible that the increased attentional processing of nodules cued by eye movement preview is what determines improved performance in this condition.

We finally note that several studies are now converging on the idea that eye movement patterns can be used as part of the process of training expertise in a variety of domains [11,12,21]. However, there remains a lack of theoretical understanding of the processes that determine the facilitatory effects of eye-movement cues in these training contexts. We would argue that a more systematic examination is required to determine what people are interpreting when they view eye movements and how, as educators, we can capitalise on the facilitating effects of eye movement training. There are already a number of parallel research threads under investigation which could extend a theoretical understanding of eye movement interpretation, such as how people make sense of real gaze behaviour in face-to-face situations [22].

5. Conclusion

The reported study demonstrates how performance in a nodule detection task can be facilitated by viewing where another person searched for nodules within the same image. We suggest that task-specific information can be obtained from viewing where another observer has searched. Further research is required to establish what types of information are accessed under these eye movement feed-forward conditions and how the interpretation of such information can be improved.

References

1. Rayner, K. (1998). Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin*, **124**, 372-422.
2. Chapman, P. R., & Underwood, G. (1998). Visual search of driving situations: Danger and experience. *Perception*, **27**, 951-964.
3. Grant, E. R., & Spivey, M. J. (2003). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science*, **14**, 462-466.
4. Pomplun, M., Ritter, H., & Velichkovsky, B. (1996). Disambiguating complex visual information: Towards communication of personal views of a scene. *Perception*, **25**, 931-948.
5. Nodine, C. F., & Kundel, H. L. (1987). Using eye movements to study visual search and to improve tumor detection. *Radiographics*, **7**, 1241-1250.
6. Nodine, C. F., Krupinski, E. A., & Kundel, K. L. (1990). A gaze-duration algorithm provides effective visual feedback to radiologists searching for nodules in chest x-ray images. *Proceedings of the First Conference on Visualization in Biomedical Computing. IEEE Computer Society Press*, 202-207.
7. 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.
8. Krupinski, E. A., Nodine, C. F., & Kundel, H. L. (1993). Perceptual enhancement of tumor targets in chest x-ray images. *Perception and Psychophysics*, **53**, 519-526.

9. Donovan, T., Manning, D. J., & Crawford, T. J. (In Press). Performance changes in lung nodule detection following perceptual feedback of eye movements. *Medical Imaging 2008 (SPIE)*.
10. 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.
11. 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. *CHI*, **36**, 289-300.
12. Stein, R., & Brennan, S. E. (2004). Another person's eye gaze as a cue in solving programming problems. *Proceedings, ICMI 2004, Sixth International Conference on Multimodal Interfaces*, 9-15.
13. Velichkovsky, B. M. (1995). Communicating attention: Gaze position transfer in cooperative problem solving. *Pragmatics and Cognition*, **3**, 199-222.
14. 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.
15. 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.
16. 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.
17. 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.
18. Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The Expertise Reversal Effect. *Educational Psychologist*, **38**, 23-31.
19. 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.
20. Chakraborty, D. P., & Berbaum, K. S. (2004). Observer studies involving detection and localization: Modeling, analysis and validation. *Medical Physics*, **31**, 2313-2330.
21. Krupinski, E. A., Tillack, A. A., Richter, L., Henderson, J. T., Bhattacharyya, A. K., Scott, K. M., Graham, A. R., Descour, M. R., Davis, J. R., & Weinstein, R. S. (2006). Eye-movement study and human performance using telepathology virtual slides. Implications for medical education and differences with experience. *Human Pathology*, **37**, 1543-1556.
22. Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze Cueing of Attention: Visual Attention, Social Cognition, and Individual Differences. *Psychological Bulletin*, **133**, 694-724.

Appendix



1. Video of a novice radiographer (green) following the eye movements of a radiologist (blue) as they search for nodules.