



Focal lung pathology detection in radiology: Is there an effect of experience on visual search behavior?

Stephanie Brams¹ · Gal Ziv² · Ignace T. C. Hooge³ · Oron Levin¹ · Thomas De Brouwere⁴ · Johny Verschakelen⁴ · Siska Dauwe¹ · A. Mark Williams⁵ · Johan Wagemans⁶ · Werner F. Helsen¹

Published online: 4 May 2020
© The Psychonomic Society, Inc. 2020

Abstract

In radiology, 60% to 80% of diagnostic errors are perceptual. The use of more efficient visual search behaviors is expected to reduce these errors. We collected eye-tracking data from participants with different levels of experience when interpreting chest X-rays during the completion of a pathology-detection task. Eye-tracking measures were assessed in the context of three existing visual search theories from the literature to understand the association between visual search behavior and underlying processes: the long-term working memory theory, the information-reduction hypothesis, and the holistic model of image perception. The most experienced participants (radiology residents) showed the highest level of performance, albeit their visual search behaviors did not differ from the intermediate group. This suggests that radiology residents better processed the represented information on the X-ray, using a visual search strategy similar to the intermediate group. Since similar visual search resulted in more information extraction in the radiology residents compared with the intermediates, we suggest that this result might support the long-term working memory theory. Furthermore, compared with novices, intermediates and radiology residents fixated longer on areas that were more important to avoid missing any pathology, which possibly confirms the information-reduction hypothesis. Finally, the larger distances between fixations observed in more experienced participants could support the holistic model of image perception. In addition, measures of generic skills were related to a lower time cost for switching between global and local information processing. Our findings suggest that the three theories may be complementary in chest X-ray interpretation. Therefore, a unified theory explaining perceptual-cognitive superiority in radiology is considered.

Keywords Radiology · Experience · Chest X-rays · Visual search behavior · Eye movements

The overwhelming majority of diagnostic errors in clinical radiology (60%–80%) are perceptual and typically arise because observers lack essential perceptual-cognitive skills

(Bruno, Walker, & Abujudeh, 2015; Kundel, Nodine, & Carmody, 1978). Perceptual-cognitive skills facilitate the individual's ability to process and integrate environmental information with existing knowledge. In radiology, at least two important perceptual-cognitive skills exist: (i) pattern recognition—specifically, similar meaningful patterns are recognized better and sooner in a specific task setting as a consequence of repetitive exposure (e.g., Loveday, Wiggins, Festa, Schell, & Twigg, 2013); and (ii) the use of situational probabilities—in this case, repetitive exposure to specific situations in a task results in early insights into the possible outcomes (Abernethy, Gill, Parks, & Packer, 2001; Hunter, 2003). Visual search measures may be linked to the underlying processes involved in perception in radiology and are monitorable or adaptable.

Indeed, Bruno et al. (2015) reported that biased visual search behavior is one of the factors responsible for perceptual errors in clinical radiology. Specifically, approximately one third of missed pulmonary nodules in clinical radiology are due to

✉ Stephanie Brams
stephanie.brams@kuleuven.be

¹ Movement Control & Neuroplasticity Research Group, Department of Movement Sciences, KU Leuven, Leuven, Belgium

² The Academic College of Wingate, Netanya, Israel

³ Experimental Psychology, Department of Psychology, Utrecht University, Utrecht, The Netherlands

⁴ Radiology, Department of Imaging and Pathology, KU Leuven, Leuven, Belgium

⁵ Department of Health, Kinesiology, and Recreation, College of Health, University of Utah, Salt Lake City, UT, USA

⁶ Laboratory of Experimental Psychology, Department of Brain & Cognition, KU Leuven, Leuven, Belgium

inefficient visual search behavior (Kundel et al., 1978). In order to identify abnormalities correctly, an observer must foveally fixate these areas (Bruno et al., 2015). Visual search can be guided by habit, practice, or previous knowledge of anatomical structures, disease patterns, and types of abnormalities (Bruno et al., 2015; Wolfe, 2012; Wolfe, Cave, & Franzel, 1989).

An individual's visual search behavior relates to how he or she directs the visual system to extract relevant information from the environment in order to make correct decisions. This visual search behavior, which differs between experts and non-experts within a specific domain of expertise (e.g., Kundel & Nodine, 1975; for a review, see Brams, Ziv, Levin, Spitz, Wagemans, Williams, & Helsen, 2019), can be described by a mixture of fixations, saccade amplitudes, distances between fixations, dwells, and so forth. Fixations, for example, keep the gaze stable on environmental stimuli, while saccades are rapid eye movements, typically from one fixation location to another (Komogortsev & Karpov, 2013). A dwell is described as one visit to an area of interest (AOI), and returns to the AOI were counted as new dwells (Holmqvist et al., 2011). So, dwell time is the duration gaze remains inside the AOI measured from entry to exit. The total dwell time was characterized as the sum of all dwell times in one AOI. Finally, entries and exits to an AOI (i.e., the number of visits to an AOI) are counted as the number of dwells. The average dwell time was calculated as the average of all the dwell times in a certain AOI during the task.

Variables such as the number and duration of fixations or dwells or length of saccades, which vary according to the visual search strategy, differ between experts and nonexperts (Brams et al., 2019).

An efficient visual search strategy is characterized by two factors—namely, the ability to guide attention towards relevant cues and the speed of rejecting distractors (Todd, Hills, & Robbins, 2012). Carrigan, Curby, Moerel, and Rich (2019) reported that expert radiologists, but not naïve participants, are cued by chest nodules, which suggests that these nodules are salient, capturing their attention. Among other factors, selective attention can be guided by information stored in long-term memory, which is gathered based on relevant experience. These memories can be retrieved relatively quickly and can strongly influence what we perceive, as proposed by long-term working memory theory (Ericsson & Kintsch, 1995). For example, expert radiologists might recognize irregularities on chest X-rays (e.g., nodules) very quickly, because their long-term memory contains many previous cases from earlier in their career. Due to this experience, essential information is represented in their working memory for rapid extraction, resulting in rapid recognition of irregularities. Due to the availability of task-related essential information in the working memory, eye-tracking data are expected to show fewer fixations of shorter duration in the X-ray image in general and irrespective of the indicated areas of interest (Gegenfurtner, Lehtinen, & Säljö, 2011).

Moreover, experts in radiology are expected to differentiate rapidly between abnormal and normal lung tissue and to know that the liver is mostly a redundant area to inspect in chest X-rays (Kok, De Bruin, Robben, & Van Merriënboer, 2012). Thus, experts are expected to be able to guide their attention to the target items relatively quickly and to quickly identify less relevant areas of the scene (Todd et al., 2012). This visual search strategy was defined in the “guided search” model (Wolfe et al., 1989) and later, from a more applied perspective, in the information-reduction hypothesis (Haider & Frensch, 1999). The guided search model or information-reduction hypothesis describes a part but not the complete experts' visual search strategy. According to the information-reduction hypothesis, dwells are expected to have longer durations, and more dwells are expected on relevant areas compared with distractor areas (Gegenfurtner et al., 2011). However, the more experienced the observer, the faster the attention will be redirected to relevant areas. Specifically, less-experienced observers (e.g., radiology residents) conduct mainly a “top-down” search that is based on their previous knowledge and the clinical question at hand (Kok et al., 2012; Wolfe & Horowitz, 2017). Although, this search is expected to be influenced more by “bottom-up” mechanism in highly experienced observers (e.g., expert radiologists), because relevant areas become salient and attract the immediate attention of experts (Carrigan, Wardle, & Rich, 2018; Carrigan et al., 2019). This guided search is further driven by “scene knowledge” (Wolfe & Horowitz, 2017) or the ability to understand where certain objects are expected to appear.

In radiology, “scene knowledge” allows experts to look less at the image than novices, because they know where to look for abnormalities in a chest X-ray (Kundel & La Follette, 1972; Kundel & Nodine, 1975). In these cases, features of the targets remain the same, but the global scene tells the expert where to look. As proposed by the holistic model of image perception (Kundel, Nodine, Conant, & Weinstein, 2007), experts are able to extract significant information about the scene with a brief glimpse (Biederman, Rabinowitz, Glass, & Stacy, 1974). This very fast extraction of the global image allows experts to access the “scene gist” (i.e., the understanding the basic meaning of the scene), which results in an efficient assessment and orientation in the complex scene (Drew, Evans, Vö, Jacobson, & Wolfe, 2013; Evans, Wolfe, Tambouret, & Wilbur, 2010; Evans, Georgian-Smith, Tambouret, Birdwell, & Wolfe, 2013; Oliva and Torralba, 2006; Treisman, 2006). During this brief glimpse, the observer is preattentively processing the scene (i.e., before intentionally attending to certain scene locations). In the next phase, gaze will be selectively directed towards targets (e.g., pathologies on chest X-rays). Eye-tracking data show a shorter time to first fixation on the relevant areas, because the participant's gaze will be attracted more strongly by abnormalities during the postattentive phase, after collecting information about the location of possible

abnormalities in the preattentive phase. Also, greater saccade amplitudes, reflecting an extended visual span, are predicted (Gegenfurtner et al., 2011). During the attentive phase, attention is moved to specific locations, and more subtle perceptual distinctions are recognized (Wolfe, Klemmen, & Dahlen, 2000; Yeshurun & Carrasco, 1999).

It is important to note that the aforementioned theories are mostly complementary rather than contradictory. However, for a given expertise domain, one theory may be more relevant than others (Brams et al., 2019). Knowledge about the contribution of each of the three theories to perceptual-cognitive superiority in a specific domain of expertise, like radiology, can provide further information concerning underlying processes explaining perceptual-cognitive skills in that domain.

Visual search strategy as indicator for experience and superior perceptual-cognitive skills in radiology

Differences in visual search behavior during medical image processing between observers with different levels of experience have been studied intensively over recent decades (see Brams et al., 2019). Overall, results strongly argue for a more global–local visual search strategy, which is in line with the holistic model of image perception, in more experienced or better performing radiologists (Bertram et al., 2016; Donovan & Litchfield, 2013; Krupinski, 1996; Mallett et al., 2014; Manning, Ethell, Donovan, & Crawford, 2006; Nodine, Kundel, Lauver, & Toto, 1996; G. Wood et al., 2013). However, a literature search analyzing the relation between gaze and expertise in medicine showed that eye-tracking analyses in medicine were often incomplete (Brams et al., 2019). Specifically, this type of research has often focused on the holistic model of image perception, which raised the question whether other theories and their related eye-movement characteristics were neglected. This gap in the literature prevents us from gaining insight to the importance of each theory while interpreting chest X-rays (Brams et al., 2019).

There is some support for each of the theories in clinical radiology. For example, shorter dwell times on nodules and less fixations to interpret chest X-rays were reported in experts for successful diagnoses (Donovan & Litchfield, 2013; Manning et al., 2006). The fact that shorter dwell times and a smaller number of fixations on nodules are observed in better performing experts might suggest that they more rapidly recognize the nodules due to information that is represented in their working memory. So, these results are in line with the long-term working memory theory (Ericsson & Kintsch, 1995). Moreover, in novices, attention is often attracted to less-relevant but salient areas, such as the stomach on chest X-rays (Kok et al., 2012). This observation is in line with the information-reduction hypothesis, which proposes that

experts, but not novices, are able to direct their attention towards relevant areas and ignore irrelevant areas (Haider & Frensch, 1999). Finally, expert radiologists appeared to fixate faster on lung nodules during interpretation of chest X-rays (Donovan & Litchfield, 2013), which is in line with the holistic model of image perception (Kundel et al., 2007). According to a previous meta-analysis, visual search is strongly dependent on the representation of images, the task, instructions, and participants (Gegenfurtner et al., 2011). As stated in the systematic review of Brams et al. (2019), generalizing eye-tracking results across different studies might result in interpretation biases about underlying mechanisms related to the three theories. It is important to collect complete sets of eye-tracking measures related to all three theories while completing one specific task (focal lung pathology-detection on chest X-rays in this study) in one domain of expertise (radiology in this study), to identify to what extent the three theories explain perceptual-cognitive expertise in that specific domain. Currently, the data are often incomplete—for example, attention (dwells) is often only measured on the area of pathology, and global scanning behavior (distance between fixations) is seldomly measured in radiology. As a consequence, it is very difficult to draw conclusions about the three theories within one specific professional task, like chest X-ray interpretation (Brams et al., 2019).

Systematic scan patterns as indicators for experience and superior perceptual-cognitive skills in radiology

Specific visual scan patterns have been identified as predictors of expert performance across multiple domains of expertise (Brams et al., 2019). However, some contradictory findings about the relationship between scan pattern systematicity and expertise can be found in the literature. For example, in a study comparing visual search in aircraft pilots with nonpilots while completing a cockpit task, no higher systematicity in visual scanning was observed in the pilots compared with the nonpilots (Brams et al., 2018). In general, the relationship between scan pattern systematicity and expertise is rarely analyzed, and it may be an important measure in various expertise domains (Brams et al., 2019).

In radiology, after being attracted to possible abnormalities as a consequence of the preattentive phase, a second and more detailed inspection of the rest of the scene will be conducted to prevent the observer missing less conspicuous abnormalities. Published reports suggest that this search follows a systematic scan pattern in expert radiologists, which is linked to better performance in interpreting chest X-rays (Augustyniak & Tadeusiewicz, 2006; Crespi, Robino, Silva, & De'Sperati, 2012; Dreiseitl, Pivec, & Binder, 2012; Kok et al., 2012;

Kok et al., 2016; Leong, Nicolaou, Emery, Darzi, & Yang, 2007; Li, Shi, Pelz, Alm, & Haake, 2016; Stockman, 2016).

We can assume that systematicity of the visual search behavior is a predictor for superior performance on a perceptual-cognitive task such as X-ray interpretation. In this study, we used an innovative measure to analyze a possible relationship between systematic scan patterns and experience—namely, the transition entropy (Allsop & Gray, 2014). For the analysis of chest X-rays, transition entropy can show whether participants scan relevant areas in a specific order. It appears that repetition of a specifically ordered scan pattern can be beneficial for task performance in radiology (Kok et al., 2016).

Generic attentional skills as indicators of experience and superior perceptual-cognitive skills in radiology

It is plausible that generic attentional skills result in better perceptual cognitive task performance. In radiology, much is reported about experts being able to extract information from an X-ray on a first glimpse, resulting in an efficient local search afterwards (Brams et al., 2019; Evans, Haygood, Cooper, Culpán, & Wolfe, 2016; Sheridan & Reingold, 2017). They are constantly switching between global and local search, as described in the holistic model of image perception (Kundel et al., 2007). This capability is tested in the Navon selective attention task (Navon, 1977). To our knowledge, this test has not previously been used in radiology, and it is difficult to make predictions about the relation between performance in this generic test and pathology-detection performance. Consequently, this is an interesting measure to analyze in groups with different levels of experience in radiology and performance on chest X-rays.

Based on our systematic literature review (Brams et al., 2019) and the most relevant lines of work regarding radiology summarized above, the purpose of the current study was to answer the following research questions: (i) Do certain theories of visual search complement one another in explaining superior perceptual-cognitive skills in chest X-rays pathology detection? (ii) Is there a difference in scan pattern systematicity between different levels of experience in radiology? (iii) Is there a specific attentional generic skill that evolves with experience in radiology that might provide more support for the theories and explain eye-tracking results?

In line with Gegenfurtner et al.'s (2011) proposed theories and related visual search behavior, we hypothesized that more experience in interpreting chest X-rays would be related to (1) shorter fixation durations and fewer eye movements, which is in line with the long-term working memory theory; (2) more dwells of a longer duration on the areas of pathology and other areas of interest, which is in line with the information-reduction hypothesis; (3) a larger distance between fixations

and a shorter latency to first dwell, which is in line with the holistic model of image perception. Besides, more experience is expected to be related to (4) the use of a more systematic scan pattern during completion of a focal lung pathology-detection task; and better performance on (5) the Navon selective attention task.

Materials and methods

Participants

Based on previous research in radiology we expected moderate to high effect sizes. Our main interest was to examine whether participants with different level of experience use different gaze characteristics when interpreting chest X-rays with or without pathologies. Using G*Power (Faul, Erdfelder, Buchner, & Lang, 2009), we calculated that a total of 42 participants is required in order to achieve 80% statistical power for an interaction of a moderate effect, and that a total of 21 participants is required for a large effect.

We recruited 41 participants. All participants had normal or corrected-to-normal visual acuity. Participants were classified into three groups. The first group (novices) consisted of medical students (2nd to 4th year of their medical education) who had already received theoretical courses about how to analyze chest X-rays ($n = 15$, mean age = 20.6 ± 1.6 years). The second group (intermediates) consisted of medical residents (seven residents in internal medicine; two junior residents in radiology; one resident in neurology; one resident in orthopedy; one resident in pneumology, and one general practice resident) with practical experience in interpreting chest X-rays (approximately 200 cases; $n = 13$, mean age = 26.8 ± 1.2 years). All participants in the intermediate group had analyzed fewer cases than any of the radiology residents. Finally, the third group consisted of radiology residents who were in their third or fourth specialization year (on average 3,000 chest X-rays cases a year; $n = 13$, mean age = $27.8 \pm .7$ years). Participants were contacted personally or recruited by the university hospital (UZ Leuven, Gasthuisberg). They all received an e-mail with additional information regarding the protocol and focus of the experiment. Participants provided written informed consent, and ethical approval was provided by the local university (KU Leuven) ethics committee (G-201504218).

Equipment and tasks

Focal lung pathology-detection task Participants were instructed to detect focal lung pathologies as accurately as possible on 26 chest X-rays (image size: $1,013.47 \times 1,023.57$ px). The chest X-rays were selected and analyzed by two professional radiologists (coauthors J.V. and T.D.). Only straightforward X-rays with focal lung diseases or

normal X-rays were included in this study (chest X-ray images with unclear cases and diffuse diseases were excluded). Each chest X-ray appeared for 60 seconds on the screen of the Tobii 1750 eye tracker (resolution: $1,280 \times 1,024$ pixels). The chest X-rays contained either no pathology ($N = 8$), one pathology ($N = 14$), two pathologies ($N = 1$), or three pathologies ($N = 3$). The focal lung pathologies were nodules and other focal consolidations (i.e., focal lung pathology in which the alveoli are filled with fluid, pus, blood or cells, instead of air) with a diameter that ranged between 0.5–5 cm, located near the diaphragm, hila, in the middle of the lung area, or in the upper lung part. Participants' eye movements were recorded throughout the task. Participants were instructed to hold their head as steady as possible, and the head was stabilized using a chin rest to minimize unintended movements. Eye-tracking data were processed using the Tobii Studio Version 3.2.1 software at a sampling frequency of 50 Hz. Raw data were extracted from the Tobii Studio Version 3.2.1 software and saved. Further processing of the raw data was conducted off-line using MATLAB scripts (coauthor I.H.). We calculated fixation locations and durations by the fixation classifier of Hooze and Camps (2013). In addition, based on the fixation durations and locations we performed an area of interest (AOI) analysis that revealed dwell times and number of dwells.

The Navon selective attention task The Navon selective attention task was conducted in order to measure the ability to selectively attend to, and switch between, global and local levels of hierarchical visual stimuli (for a detailed description of the task, see Chamberlain, Van der Hallen, Huygelier, Van de Cruys, & Wagemans, 2017; Navon, 1977). On each trial, a large shape (global level) made up of 18 smaller shapes (local level) appeared. Participants were asked to indicate squares (pressing the “F” key) or circles (pressing the “J” key) quickly and accurately. These squares or circles were present either at the global or local level. Participants were not cued about whether they should search for the information at global or local level. They had to check both levels as quickly as possible. Thirty-two trial pairs were presented at random. One trial pair consisted of two trials in which the circle or square had to be detected in one trial, either at a global (G) or a local (L) level, so the possible pair combinations were: GL; LG; GG; LL. Reaction times could be influenced by the type of trial pair that occurred, especially when a switch from local to global detection, or vice versa, was necessary. Mean global–local and local–global reaction time costs were computed based on the subtractions of local from global (GL pair) and global from local (LG pair) trial reaction times. During the Navon selective attention task, accuracy as well as reaction time were measured. Accuracy was calculated separately for global and local detection, as a percentage of number of correct trials.

Reaction time was measured and used as an indication of speed in attentional switching. For this analysis, the difference in reaction time (RT) to detect a global figure followed by a local figure and vice versa was calculated, which indicates the time cost to switch attention from global to local detection ($GL = RT_{Global} - RT_{Local}$) and from local to global detection ($LG = RT_{Local} - RT_{Global}$). During each generic task, feedback was given by a green cross for correct responses or a red cross for incorrect responses.

Procedure

Before starting the focal lung pathology-detection task, five X-rays were shown in order to familiarize participants with the screen resolution and the X-rays. The five X-rays used during the familiarization phase were excluded from the main experiment. The participants were seated at a distance of 60 cm from the Tobii monitor and placed their chin in a chin rest. Next, a 9-point gaze calibration was performed. The participants were instructed to hold their head steady in the chin rest throughout the entire task. Then, the participants evaluated 26 chest X-rays for focal lung pathologies, which appeared individually. Before the presentation of each image, a black screen with the trial number presented on it appeared. Participants were instructed to detect pathologies as accurately as possible and to indicate them with a mouse click. As long as the mouse click was within the borders of the area of pathology, as indicated as an area of interest (AOI) in Fig. 1a, it was considered as a correct detection.

After each image, participants were asked to respond to a multiple-choice question: “I did not click because . . .” The possible answers were: “There was no pathology”; “I was too late”; “I have no idea”; or “irrelevant.” If participants did not identify a pathology, the question allowed them to explain why by clicking one of the first three options. If participants did identify a focal lung pathology, they could respond with the last option (“irrelevant”). The entire session lasted approximately 30 minutes. No feedback was given to the participants, and gaze behavior was recorded throughout the task.

After completing the focal lung pathology-detection task, the Navon selective attention task was completed. This task was completed on a laptop computer and took approximately 15 minutes. Eye movements were not recorded.

Dependent variables and data processing

All data are available at the following link: <https://osf.io/pz3sc/>

Focal lung pathology-detection task performance A d' -prime (d') value was calculated based on the average number of correctly identified X-rays with pathology and normal X-rays. Specifically:

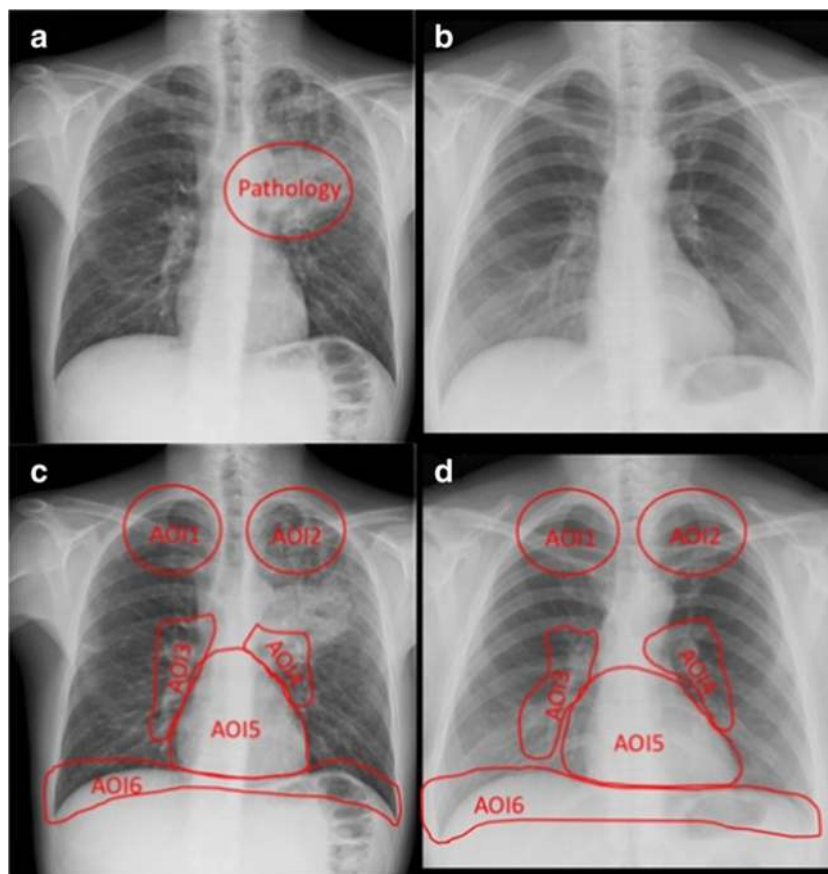


Fig. 1 Representation of chest X-rays used in the focal lung pathology-detection task with AOIs. **a, c** X-ray with pathology, (**a**) focal lung pathology indicated as an AOI and (**c**) other AOIs. **b, d** Thorax without

pathology (**b**), indicated AOIs (**d**). AOI1 = right upper lung; AOI2 = left upper lung; AOI3 = right hila; AOI4 = left hila; AOI5 = heart area; AOI6 = diaphragm

$$d' = Z(\text{Hit rate}) - Z(\text{False alarm rate})$$

$$\text{Hit rate} = \frac{\text{Correct identified pathologies}}{\text{Total number of pathologies}} \quad \text{False alarm rate} = \frac{\text{normal Xrays that were clicked on}}{\text{Total number of normal Xrays}} \quad (1)$$

Response time was calculated as the time from the onset of the trial until the last mouse click. That is, for each chest X-ray, response time refers to the time from the onset of the trial until the participant decided that all pathologies were detected. For each participant, these values were averaged over all trials.

Eye-tracking data All eye-tracking data were assessed from the onset of the trial until the participant's last mouse click. This was done as most participants, especially novices, tended to relax and stopped monitoring the chest X-ray after their last mouse click. Data recording prior to the last mouse-click lasted on average 16.6 ± 5.0 seconds.

Visual search strategies Each focal lung pathology was indicated as an AOI. AOIs had an average area size of $11,735.8 \pm 7,840.8$ pixels (see Fig. 1a–b). Six other AOIs were indicated

by an expert radiologist with more than 20 years of experience in clinical radiology: left ($12,150.5 \pm 2,530.3$) and right upper lung areas ($12,887.4 \pm 2,867.1$); left ($4,804.7 \pm 1,746.7$) and right hila area ($7,112.9 \pm 2,124.2$); heart area ($20,989.2 \pm 4,283.4$); and diaphragm area ($25,532.7 \pm 7,252.2$; see Fig. 1c–d). The area outside the AOI's is defined as “rest” area and contains less relevant (healthy, noncrowded lung tissue) or irrelevant information (liver, bowels, and arms) for chest X-ray patient analysis.

Visual search strategies were assessed with respect to (1) the whole scene irrespective of AOIs: average fixation duration and average number of eye movements (in line with the long-term working memory theory), average fixation distance (in pixels; Over, Hooge, & Erkelens, 2006; in line with the holistic model of image processing), and (2) the AOIs: entropy (to assess scan pattern systematicity), average latency to first

dwell on the area of pathology (in line with the holistic model of image processing), average number of dwells, average dwell time, and number of missed AOIs throughout the entire task (in line with the information-reduction hypothesis).

The average fixation duration, average fixation distance, and average number of eye movements were calculated separately for the chest X-ray with pathologies and for the chest X-ray without pathologies. Differences between both measures were analyzed to examine changes in search behavior on chest X-rays that contain a focal lung pathology versus chest X-rays with no pathology. We used entropy to indicate systematicity of the scan pattern, which was calculated according to the method of Allsop and Gray (2014). The method is based on a transition matrix indicating the chance that the participant's fixation will move from one specific AOI to another specific AOI in a particular order (see Allsop & Gray, 2014; Brams et al., 2018).

The average fixation distance as well as the latency to the first dwell on the area of pathology (e.g., the time interval from the time that an X-ray was shown until the first dwell on the AOI) provided measures to support holistic image processing. A wide fixation distance (representing a wide visual span) suggests global scene processing, which is expected to efficiently guide attention afterwards during the detailed local search. As a consequence, a shorter latency to first dwell is expected. A dwell was defined as one visit to an AOI, and returns to the AOI were counted as new dwells (Holmqvist et al., 2011). Dwell time is the duration gaze remains inside the AOI measured from entry to exit. The total dwell time was characterized as the sum of all dwell times in one AOI. The number of dwells was defined as the number of visits to an AOI. The average dwell time was calculated as the average of all the dwell times in a certain AOI during the task. The number of not fixated AOIs over the whole task is the sum of all pathologies that were not fixated during task completion.

Data analysis

Performance on the focal lung pathology-detection task was assessed calculating a d' -prime value for each group of participants as indicated in the methodology. A one-way analysis of variance (ANOVA) was conducted to assess group differences in response time. For the generic tasks, outcome measures were analyzed using the nonparametric statistics (Kruskal–Wallis test) because after log-transformation the values remained skewed.

Eye-tracking measures that were examined irrespective of AOI (average fixation duration, average fixation distance and average number of eye movements) were analyzed using a two-way ANOVA (Group \times Image) with repeated measures on the Image factor. The average number of eye movements were first log-transformed to obtain normality. Comparison between X-ray types (normal or focal lung pathology) was

conducted for these global measures irrespective of AOI because, based on literature, global scene processing is supposed to differ when a focal lung pathology is present (Bertram, Helle, Kaakinen, & Svedstrom, 2013; Bertram et al., 2016). When significant main effects or interactions were found, Bonferroni post hoc analyses were conducted to identify specific differences in visual search characteristics as function of X-ray types and groups. Finally, two stepwise multiple regression analyses were conducted to examine (i) the association between search behaviors on X-rays with pathology and detection of those pathologies (true positives), and (ii) the association between search behaviors on normal X-rays and correct identification of those normal X-rays (true negatives).

Other eye-tracking data were analyzed regarding specific AOIs (entropy, average latency of the first dwell on the area of pathology, number of not fixated pathologies, average number of dwells, and average dwell duration). Data were assessed using one-way ANOVAs. These analyses were conducted on three different types of AOIs, separately, to avoid overlaps. Specifically, one analysis was conducted over the areas of pathology, one over other AOIs (upper lung parts, hila, heart region, and diaphragm) and one over the “rest” area (see description above). Data for the number of dwells were first log-transformed to obtain normality. Bonferroni post hoc analyses were performed to examine specific group differences in eye-tracking measures. Finally, to control for multiple comparisons for the main effects, the false discovery rate (FDR) method was used (Benjamini & Hochberg, 1995).

Data exclusion

Trials with more than 20% of eye-movement data loss were excluded, resulting in the exclusion of 2.9% of all trials (Brams et al., 2018). Furthermore, values larger or smaller than the average plus or minus three standard deviations were excluded as outliers, resulting in 1.2% of all data being excluded (Brams et al., 2018).

Results

Performance scores

Performance on the focal lung pathology-detection task A d' -prime calculation was used as a sensitivity index for pathology detection. The results of this analysis showed that radiology residents were able to detect focal lung pathologies and normal X-rays with an accuracy above chance level ($d' = 2.62$). However, the intermediates and novices were not able to do so and performed at around chance level (Intermediates: $d' = -.20$; Novices: $d' = -1.06$).

In addition, a one-way ANOVA for response time showed significant group differences, $F(2, 40) = 15.81$, $p < .001$, $\eta_p^2 =$

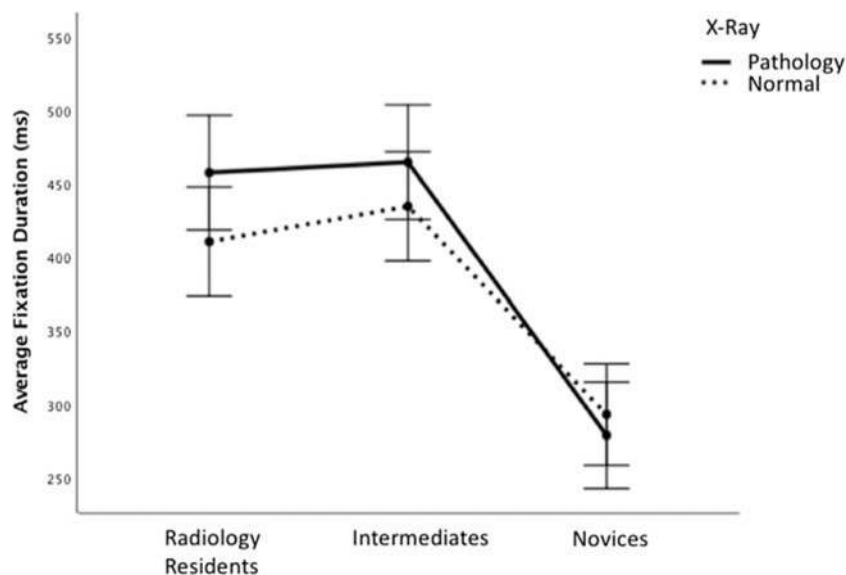


Fig. 2 Average fixation duration on chest X-rays per group

.44. A Bonferroni post hoc analysis showed that both radiology residents (13.8 ± 1.1 s) and intermediates (14.3 ± 1.1 s) had shorter response times compared with novices (21.1 ± 1.0 s).

Performance on the Navon selective attention task Most of the generic performance measures (accuracy local, accuracy global, LG and GL) were not distributed normally, even after log-transformation. For this reason, all performance measures were analyzed using the nonparametric Kruskal–Wallis test. The results showed only a significant group effect for GL, $F(2, 39) = 6.42$, $p = .04$, $\eta_p^2 = .12$ —that is, the time cost to switch from global to local information processing. Pairwise comparisons indicated that radiology residents (time cost for switch = 54.7 ± 127.5 ms) and intermediates (time cost for switch = 33.6 ± 139.2 ms) were significantly faster in switching from global to local information processing compared with novices (time cost for switch = 328.6 ± 695.5 ms).

Eye-tracking measures

Average fixation duration We used a two-way repeated-measures ANOVA (Group \times Image) for the average fixation duration to assess whether visual search differed depending on the presence of a focal lung pathology on the shown image. The results show a significant interaction between group and image, $F(2, 38) = 8.42$, $p = .001$, $\eta_p^2 = .31$. Both radiology residents' and intermediates' average fixation duration was faster when the shown image was normal (438.1 ± 140.9 and 439.9 ± 133.4 ms, respectively) compared with when a focal lung pathology was present (487.2 ± 157.8 and 469.2 ± 128.8 ms, respectively), whereas in novices the average fixation duration did not change with the presence of pathologies (292.5 ± 96.5 and 275.8 ± 86.4 ms; see Fig. 2). Also, a significant effect for group was observed, $F(2, 38) = 9.53$, $p < .001$, $\eta_p^2 = .33$. A Bonferroni post hoc analyses showed that

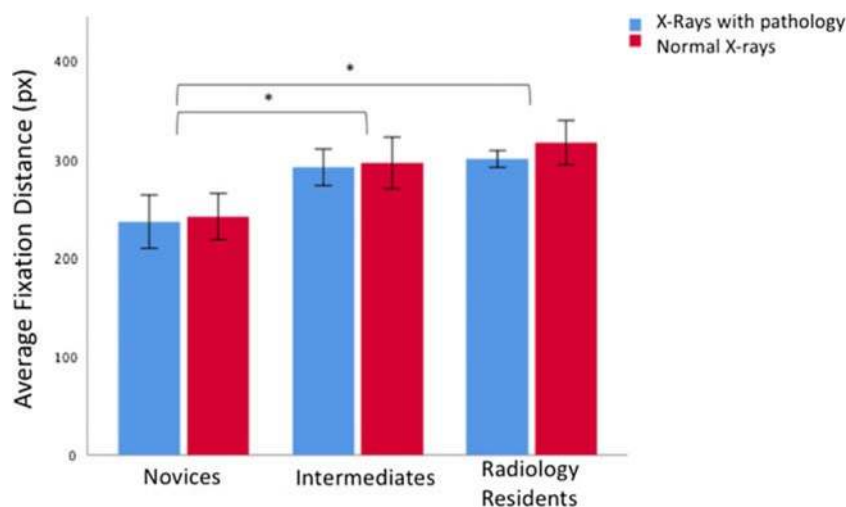


Fig. 3 Average fixation distance on chest X-rays per group

both radiology residents (458.7 ± 34.3 ms) and intermediates (457.2 ± 34.3 ms) used significantly longer fixations compared with novices (282.0 ± 31.9 ms).

Average fixation distance We employed a two-way ANOVA (Group \times Type of Image) with repeated measures on the type of image factor (with or without a pathology) to assess both differences in average fixation distance between groups. The results showed significant main effects for type of image, $F(1, 38) = 12.94, p = .001, \eta_p^2 = .25$, and group, $F(2, 38) = 16.06, p < .001, \eta_p^2 = .46$. Pairwise comparisons showed significantly greater average distance between fixations in radiology residents (298.1 ± 27.9 px) and intermediates (294.3 ± 27.9 px) compared with novices (245.0 ± 27.9 px) (21.6% and 20.1% wider in residents and intermediates, respectively; see Fig. 3). Also, a greater average fixation distance when a normal X-ray (282.3 ± 35.3 px) was analyzed compared with a pathological X-ray (272.7 ± 35.3 px) (3.5% wider in normal X-rays). No significant effect was observed for the Group \times Image interaction, $F(2, 38) = 1.88, p = .17, \eta_p^2 = .09$.

Eye-tracking measures for areas of pathology and other AOIs

Entropy analysis over AOIs We conducted a one-way ANOVA on the entropy values. The results showed no significant group effects, $F(2, 40) = 2.07, p = .14, \eta_p^2 = .10$.

Average latency of the first dwell on the area of pathology A one-way ANOVA showed no significant group differences in average latency to first dwell on the area of pathology, $F(2, 39) = 2.52, p = .09, \eta_p^2 = .12$; novices: 8.8 ± 3.6 s; intermediates: 7.7 ± 3.4 s; radiology residents: 6.4 ± 2.6 s.

Number of not fixated pathologies No significant group differences in the number of pathologies that were not fixated were observed after conducting a one-way ANOVA, $F(2, 40) = 2.82, p = .07, \eta_p^2 = .12$; novices: 5.6 ± 2.3 ; intermediates: 7.3 ± 1.8 ; radiology residents: 4.9 ± 2.8 .

Average number of dwells We conducted a log transformation on the data due to the skewed distribution of the number of dwells. A one-way ANOVA on this log-transformed data showed significant differences between groups for the average number of dwells on the area of pathology, $F(2, 40) = 9.15, p = .001, \eta_p^2 = .31$, the rest area (noncrowded healthy lung tissue and irrelevant areas like arms and liver), $F(2, 40) = 14.05, p < .001, \eta_p^2 = .43$, as well as on the other AOIs, $F(2, 40) = 11.21, p < .001, \eta_p^2 = .40$. A Bonferroni post hoc analysis showed that the average number of dwells on the areas of pathology, when interpreting chest X-rays was lower in radiology residents ($2.2 \pm .5$) and intermediates ($2.3 \pm .3$) compared with the novices (3.9 ± 1.6). For the average

number of dwells on the other AOIs, a Bonferroni post hoc analysis showed that radiology residents ($3.0 \pm .5$) and intermediates ($2.8 \pm .5$) visited these areas significantly less than novices (4.3 ± 1.3). Finally, a Bonferroni post hoc analysis showed that novices visited the rest area almost twice as much (11.9 ± 4.5) compared with intermediates (6.9 ± 1.5) and radiology residents (6.9 ± 1.7).

Average dwell duration A one-way ANOVA of the average dwell duration on the area of pathology showed no significant group difference, $F(2, 40) = .52, p = .60, \eta_p^2 = .03$; novices: $2.3 \pm .8$ s; intermediates: $2.1 \pm .4$ s; radiology residents: $2.1 \pm .5$ s.

A similar analysis over the other AOIs showed a significant group effect, $F(2, 40) = 4.77; p = .01; \eta_p^2 = .20$. A Bonferroni post hoc analysis indicated significant longer dwells in radiology residents ($1.07 \pm .29$ s) and intermediates ($1.0 \pm .3$ s), compared with novices ($.8 \pm .2$ s).

Relationship between visual search behavior and performance

A stepwise multiple regression analysis assessing the relation between visual search behavior (fixation duration and fixation distance) while inspecting X-rays with pathology and the number of true positives indicates that 18% in the variance of the true positives can be explained by the fixation distance ($p = .006$).

A similar analysis assessing the relation between visual search behavior while inspecting normal X-rays and the number of true negatives showed that 43% of the variance in the correct identified true negatives could be explained by the fixation distance ($p < .001$).

In general, the stepwise regression analyses indicated that a larger fixation distance might be a predictor for a higher accuracy on the focal lung pathology-detection task.

Discussion

We examined the effects of experience on chest X-ray focal lung pathology detection via differences in visual search strategies. In line with Gegenfurtner et al.'s (2011) proposed theories related to visual search behavior, we hypothesized that more experience in interpreting chest X-rays would be related to (i) shorter fixation durations and fewer eye movements; (ii) more dwells of a longer duration on the areas of pathology and other areas of interest; (iii) a larger distance between fixations and a shorter latency to first dwell. Also, more experience in interpreting chest X-rays is expected to be related to (iv) the use of a more systematic scan pattern during completion of a focal lung pathology-detection task, and (v) better performance on the Navon selective attention task.

Experience level was related to focal lung pathology-detection performance. Specifically, based on *d*-prime calculations, radiology residents were able to detect focal lung pathologies and normal chest X-rays with an accuracy above chance level. In contrast, the performance of intermediate and novice participants was at chance level.

Furthermore, both radiology residents and intermediates (i.e., all participants with practical experience in interpreting chest X-rays) had faster response times, and longer average fixation durations, compared with novices. Finally, we showed that there is a specific attentional generic skill that evolves with experience in radiology. The radiology residents and intermediates had faster global to local processing in the Navon selective attention task compared with novices.

To better understand which underlying processes result in higher detection performance, we addressed three theories of visual search strategies and expert performance: (1) the long-term working memory theory (Ericsson & Kintsch, 1995); (2) the information-reduction hypothesis (Haider & Frensch, 1999); and (3) the holistic model of image perception (Kundel et al., 2007). In addition, we examined the relation between level of experience and scan systematicity. Previously, researchers have shown that more systematic scanning might be a predictor of superior perceptual-cognitive skills and experience (Allsop & Gray, 2014; Kok et al., 2016; O'Neill et al., 2011; Ziv, 2017).

Visual search strategy as indicator of experience and superior perceptual-cognitive skills in radiology

Our first hypothesis was that the group with the highest level of experience (radiology residents) will use distinctive visual search strategies that are in line with the three proposed theories of perceptual-cognitive superiority (Gegenfurtner et al., 2011).

The long-term working memory theory The observation that radiology residents had faster response times, whereas their latency to first dwell the areas of pathology did not differ compared with the other groups, providing support for the long-term working memory theory. Specifically, compared with other groups, the residents were able to decide whether a pathology was present more rapidly, even though the time to first fixation of the pathology was similar to the other groups. Furthermore, both radiology residents and intermediates used longer fixation durations irrespective of AOI compared with novices. This finding indicates that practitioners with practical experience in interpreting chest X-rays use a slower visual search rate (i.e., longer fixation durations, smaller number of dwells, and longer dwell durations on AOI's) compared with those with only a theoretical background about chest X-rays. This observation contradicts the long-term working memory theory, since a slower visual search rate represents longer

fixation durations, whereas, according to this theory, shorter fixation duration would be expected (Gegenfurtner et al., 2011). The results of the current study are supported, nonetheless, by a previous meta-analysis which showed that a slow visual search rate might be a more efficient visual search strategy in detection tasks (Gegenfurtner et al., 2011). Furthermore, this result is in line with a recent systematic review reporting that a slow visual search rate seems to be the strategy that is mostly used in experts in medical image processing (Brams et al., 2019). A possible explanation for this observation might be that it is more efficient to use a visual search strategy that is more controlled (i.e., spending longer on more relevant areas) and less exhaustive (i.e., using longer fixations and fewer dwells; Donovan & Litchfield, 2013; Milazzo, Farrow, Ruffault, & Fournier, 2016). This does not necessarily mean that the time to complete the task will be longer.

However, visual search rate was faster when the X-ray was normal compared with when focal lung pathology was present. Specifically, radiology residents and intermediates used shorter fixation durations when the chest X-ray was normal compared with when focal lung pathology was present, resulting in a faster visual search rate, whereas the average fixation duration did not change in novices. It is possible that due to prior knowledge, an X-ray type (normal or with pathology) is easily recognized in the groups with practical experience. In the case of a normal X-ray, a more rapid check (shorter fixation durations) can then be justified, whereas when pathologies are present, a more detailed processing of information at the area of interest (i.e., scrutiny of possible targets) is employed in order to not miss any focal lung pathology (resulting in longer fixation durations). This adaptive behavior partly supports the long-term working memory theory.

Our results are in line with previous published reports in which experts, intermediates, and novices performed a detection task on CT scans searching for lesions, enlarged lymph nodes (ELNs), and/or visceral abnormalities (Bertram et al., 2013; Bertram et al., 2016). The experts in these studies used longer fixation durations when ELNs or lesions were present compared with CT scans without ELNs or lesions. This adaptive behavior was not observed in the intermediates and novices in this study. Moreover, using longer fixation durations in the presence of focal lung pathology is expected to result in higher task accuracy, because longer fixation durations allow for better identification (Bertram et al., 2016).

The information-reduction hypothesis A significantly lower number of dwells in the area of pathology, the rest area (noncrowded healthy lung tissue, arms, bowels, and liver) and in other anatomical AOIs (upper lung parts, hila, heart region, and diaphragm) was observed in the two groups with practical experience, compared with the novice group. A large

effect size was obtained for the number of not fixated pathologies suggests that radiology residents fixate on more areas of pathology compared with other participants.

Furthermore, dwell duration in the area of pathology did not differ between groups, whereas dwell duration in other AOIs was significantly longer in the two groups with practical experience compared with novices. In chest X-rays, these AOIs are more crowded, so these are the areas where pathologies are often missed during inspection of chest X-rays and therefore require more detailed attention (Swenson, Hessel, & Herman, 1977). Furthermore, the areas of the bowels, liver, and healthy noncrowded lung tissues are assumed to be less relevant areas (Kok et al., 2012). Our results showed that novices visited these areas twice as much compared with intermediates and radiology residents. The results partly support the information-reduction hypothesis, since this hypothesis argues for a higher attention allocation towards important informative areas and less attention allocation towards less relevant or irrelevant areas (Haider & Frensch, 1999). Finally, the observation that radiology residents fixate a higher number of the presented pathologies further supports the information-reduction hypothesis.

The holistic model of image perception The lower number of dwells on the areas of pathology and other anatomical AOIs observed in the two more experienced groups supports the holistic model of image perception (Kundel et al., 2007). According to this model, the development of an extended visual span, such as the enhanced capability to process information parafoveally (Rosenholtz, 2016), in more experienced participants allows them to process information during a global search. Thereafter, relevant areas are “double checked” during local searches (Leong et al., 2007).

Furthermore, radiology residents and intermediates showed longer average distance between fixations, compared with novices. This finding suggests a wider visual span in the two more experienced groups compared with the novice group, which also supports this model. The holistic model of image perception can explain the decreased visual span with the presence of focal lung pathology. Based on this model, it is expected that areas of pathology will be fixated faster from a first global glimpse. This will be followed by shorter saccades in a local search. This observation is in line with those reported in a previous study in which shorter saccade amplitudes were observed in experts when lesions were present on CT images compared with normal CT images (Bertram et al., 2016).

Finally, results of a stepwise regression analysis showed that greater distances between fixations, while inspecting both normal X-rays and X-rays with pathology, was related with a higher detection rate of true negatives and true positives. So, a large distance between fixations is expected to be advantageous for task performance as it allows to capture the

complete global scene before a local search is conducted. This, in turn, may lead to a more efficient local search (Leong et al., 2007; Manning et al., 2006).

Systematic scan patterns as indicators of experience and superior perceptual-cognitive skills in radiology

Our data do not support the fourth hypothesis, as entropy calculations indicated that there were no differences in scan pattern systematicity between groups. Our results are aligned with those reported in previous work that examined the scan patterns of aircraft pilots and nonpilots (Brams et al., 2018). A possible explanation for not finding a relationship between systematic scanning behavior and experience is the smaller size of the Tobii screen compared with the screens used in the radiology department, which might facilitate parafoveal information processing. In this case, a structured scan might not be necessary (Brams et al., 2018). However, previous studies reported no effects of screen size on radiologists' performance (Gur et al., 2006; Mc Laughlin et al., 2012). It is also possible that the entropy measure may not be sensitive enough to assess systematicity in chest X-ray interpretation.

Finally, a recent review of 22 eye-tracking studies in radiology suggest that certain scan characteristics, such as systematic scanning of chest X-rays, are related to high levels of expertise (Van der Gijp et al., 2017). It is possible that systematic scanning only evolves after more years of experience.

Generic attentional skills as indicators of experience and superior perceptual-cognitive skills in radiology

For the Navon selective attention task results indicated, as expected, that more experienced participants (radiology residents and intermediates) were faster in switching their attention between global and local information processing. This finding suggests that they use a global–local information processing strategy, in line with the holistic model of image perception (Kundel et al., 2007) and previous research on pilots (Brams et al., 2018).

During a global–local information processing strategy, the scene is first captured globally during the preattentive phase (Kundel & Nodine, 1975). Then, the gaze is rapidly guided to the AOI for a local scan. The results of the current study suggest that this strategy is used by participants that are more experienced in interpreting chest X-rays (Leong et al., 2007; Manning et al., 2006). The global–local search strategy used in medical imaging is often called a discovery-reflective visual search. The participants are able to process the most essential information with one glance (discovery phase) and then conduct a local scan as confirmation (reflective phase). This strategy appears to result in higher detection accuracy (Leong et al., 2007; O'Regan & Levy-Schoen, 1987).

Our results suggest that the three theories proposed to explain perceptual-cognitive skills may be complementary in chest X-ray interpretation. Our radiology residents showed fewer dwells on the area of pathology (in line with the long-term working memory theory), longer dwells on other AOI's where pathologies are often missed (in line with the information-reduction hypothesis), and larger distances between fixations (in line with holistic model of image perception). However, contrary to what was expected, the most experienced residents did not show more organized, systematic, visual scanning behavior while inspecting chest X-rays.

Conclusions

We examined the effects of eye movements and visual scanning on focal lung pathology-detection performance in radiology. The results showed differences in eye movements between different levels of experience in interpreting chest X-rays. Specifically, in contrast to less experienced students (intermediates and novices), radiology residents were able to detect pathologies and normal X-rays with an accuracy above chance level. Their performance was associated with faster response times, longer average fixation durations, and faster global to local visual processing. While this study showed differences in gaze behavior and accuracy between more experienced and less experienced participants, it remains to be seen whether there is a causal relationship between more efficient gaze strategies and higher focal lung pathology-detection performance. Therefore, in future, researchers should include thoracic radiologists to examine if, and if so, to what extent, our results further evolve with increasing experience and expertise. Such information can be used to compose gaze-training programs that might lead to improved performance in radiology (Auffermann, Little, & Tridandapani, 2015; Gegenfurtner, Lehtinen, Jarodzka, & Säljö, 2017; Kok et al., 2016; van Geel, Kok, Dijkstra, Robben, & van Merriënboer, 2017).

Acknowledgements We thank the Department of Radiology at the University Hospital UZ Leuven (Gasthuisberg) for helping with the recruitment of participants and providing a room where we could conduct the experiments. For technical help with the eye tracker, we would like to thank Kristof Dockx.

Statement of significance A large proportion of errors in detecting pathologies in chest X-rays are due to inefficient visual search. In order to find a pathology on a chest X-ray, gaze needs to be used optimally. We assess the gaze characteristics that differentiate experienced and less experienced participants when interpreting chest X-rays and whether these gaze characteristics are associated with pathology detection.

Understanding the visual search strategies associated with improved pathology detection can help in the training of radiologists and reduce the number of errors in pathology detection.

References

- Abernethy, B., Gill, D. P., Parks, S. L., & Parcker, S. T. (2001). Expertise and the perception of kinematic and situational probability information. *Perception*, 30, 233–252. doi:<https://doi.org/10.1068/p2872>
- Allsop, J., & Gray, R. (2014). Flying under pressure: Effects of anxiety on attention and gaze behavior in aviation. *Journal of Applied Research in Memory and Cognition*, 3, 63–71. doi:<https://doi.org/10.1016/j.jarmac.2014.04.010>
- Auffermann, W. F., Little, B. P., & Tridandapani, S. (2015). Teaching search patterns to medical trainees in an educational laboratory to improve perception of pulmonary nodules. *Journal of Medical Imaging*, 3, 011006. <https://doi.org/10.1117/1.JMI.3.1.011006>
- Augustyniak, P., & Tadeusiewicz, R. (2006). Assessment of electrocardiogram visual interpretation strategy based on scanpath analysis. *Physiological Measurement*, 27, 597–608. doi:<https://doi.org/10.1088/0967-3334/27/7/004>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: Series B (Methodological)*, 57, 289–300. doi:<https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Bertram, R., Helle, L., Kaakinen, J. K., & Svedstrom, E. (2013). The effect of expertise on eye movement behaviour in medical image perception. *PLOS ONE*, 8, e66169. doi:<https://doi.org/10.1371/journal.pone.0066169>
- Bertram, R., Kaakinen, J., Bensch, F., Helle, L., Lantto, E., Niemi, P., & Lundbom, N. (2016). Eye movements of radiologists reflect expertise in CT study interpretation: A potential tool to measure resident development. *Radiology*, 281, 805–815. doi:<https://doi.org/10.1148/radiol.2016151255>
- Biederman, I., Rabinowitz, J. C., Glass, A. L., & Stacy, E. W. (1974). On the information extracted from a glance at a scene. *Journal of Experimental Psychology*, 103, 597. doi:<https://doi.org/10.1037/h0037158>
- Brams, S., Hooge, I. T. C., Ziv, G., Dauwe, S., Evens, K., De Wolf, T., ... Helsen, W. F. (2018). Does effective gaze behavior lead to enhanced performance in a complex error-detection cockpit task? *PLOS ONE*, 13, e0207439. doi:<https://doi.org/10.1371/journal.pone.0207439>
- Brams, S., Ziv, G., Levin, O., Spitz, J., Wagemans, J., Williams, A. M., & Helsen, W. F. (2019). The relationship between gaze behavior, expertise, and performance: A systematic review. *Psychological Bulletin*, 145, 980–1027. doi:<https://doi.org/10.1037/bul0000207>
- Bruno, M. A., Walker, E. A., & Abujudeh, H. H. (2015). Understanding and confronting our mistakes: The epidemiology of error in radiology and strategies for error reduction. *Radiographics*, 35, 1668–1676. doi:<https://doi.org/10.1148/rg.2015150023>
- Carrigan, A. J., Curby, K. M., Moerel, D., & Rich, A. N. (2019). Exploring the effect of context and expertise on attention: Is attention shifted by information in medical images? *Attention, Perception, & Psychophysics*, 1–14.
- Carrigan, A. J., Wardle, S. G., & Rich, A. N. (2018). Finding cancer in mammograms: If you know it's there, do you know where? *Cognitive Research: Principles and Implications*, 3, 10.
- Chamberlain, R., Van der Hallen, R., Huygelier, H., Van de Cruys, S., & Wagemans, J. (2017). Local-global processing bias is not a unitary individual difference in visual processing. *Vision Research*, 141, 247–257. doi:<https://doi.org/10.1016/j.visres.2017.01.008>
- Crespi, S., Robino, C., Silva, O., & De'Sperati, C. (2012). Spotting expertise in the eyes: Billiards knowledge as revealed by gaze shifts in

- a dynamic visual prediction task. *Journal of Vision*, 12. doi:<https://doi.org/10.1167/12.11.30>
- Donovan, T., & Litchfield, D. (2013). Looking for cancer: Expertise related differences in searching and decision making. *Applied Cognitive Psychology*, 27, 43–49. doi:<https://doi.org/10.1002/acp.2869>
- Dreiseitl, S., Pivec, M., & Binder, M. (2012). Differences in examination characteristics of pigmented skin lesions: Results of an eye tracking study. *Artificial Intelligence in Medicine*, 54, 201–205. doi:<https://doi.org/10.1016/j.artmed.2011.11.004>
- Drew, T., Evans, K., Vö, M. L. H., Jacobson, F. L., & Wolfe, J. M. (2013). Informatics in radiology: What can you see in a single glance and how might this guide visual search in medical images? *Radiographics*, 33, 263–274. doi:<https://doi.org/10.1148/rg.331125023>
- Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102, 211. doi:<https://doi.org/10.1037/0033-295X.102.2.211>
- Evans, K., Wolfe, J., Tambouret, R., & Wilbur, D. (2010). In the blink of an eye: Discrimination and localization of abnormalities in cervical cytology screening from a global signal. *Cancer Cytopathology*, 118, 389. doi:<https://doi.org/10.1002/cncy.20104>
- Evans, K. K., Georgian-Smith, D., Tambouret, R., Birdwell, R. L., & Wolfe, J. M. (2013). The gist of the abnormal: above-chance medical decision making in the blink of an eye. *Psychonomic Bulletin & Review*, 20, 1170–1175. doi:<https://doi.org/10.3758/s13423-013-0459-3>
- Evans, K. K., Haygood, T. M., Cooper, J., Culpan, A. M., & Wolfe, J. M. (2016). A half-second glimpse often lets radiologists identify breast cancer cases even when viewing the mammogram of the opposite breast. *Proceedings of the National Academy of Sciences of the United States of America*, 113, 10292–10297. doi:<https://doi.org/10.1073/pnas.1606187113>
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149–1160. doi:<https://doi.org/10.3758/BRM.41.4.1149>
- Gegenfurtner, A., Lehtinen, E., & Säljö, R. (2011). Expertise differences in the comprehension of visualizations: A meta-analysis of eye-tracking research in professional domains. *Educational Psychology Review*, 23, 523–552.
- Gegenfurtner, A., Lehtinen, E., Jarodzka, H., & Säljö, R. (2017). Effects of eye movement modeling examples on adaptive expertise in medical image diagnosis. *Computers & Education*, 113, 212–225. doi:<https://doi.org/10.1016/j.compedu.2017.06.001>
- Gur, D., Klym, A. H., King, J. L., Maitz, G. S., Mello-Thoms, C., Rockette, H. E., & Thaete, F. L. (2006). The effect of image display size on observer performance: an assessment of variance components. *Academic Radiology*, 13, 409–413. doi:<https://doi.org/10.1016/j.acra.2005.11.033>
- Haider, H., & Frensch, P. A. (1999). Eye movement during skill acquisition: More evidence for the information-reduction hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 172–190.
- Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). Numerosity measures. In K. Holmqvist & R. Andersson (Eds.), *Eye tracking: A comprehensive guide to methods and measures* (pp. 412–419). Oxford, England: Oxford University Press.
- Hooge, I. T., & Camps, G. (2013). Scan path entropy and arrow plots: Capturing scanning behavior of multiple observers. *Frontiers in Psychology*, 4, 996–1006. doi:<https://doi.org/10.3389/fpsyg.2013.00996>
- Hunter, D. R. (2003). Measuring general aviation pilot judgment using a situational judgment technique. *International Journal of Aviation Psychology*, 13, 373–386.
- Kok, E. M., De Bruin, A. B., Robben, S. G., & Van Merriënboer, J. J. (2012). Looking in the same manner but seeing it differently: Bottom-up and expertise effects in radiology. *Applied Cognitive Psychology*, 26, 854–862. doi:<https://doi.org/10.1002/acp.2886>
- Kok, E. M., Jarodzka, H., de Bruin, A. B., BinAmir, H. A., Robben, S. G., & van Merriënboer, J. J. (2016). Systematic viewing in radiology: seeing more, missing less? *Advances in Health Sciences Education*, 21, 189–205. doi:<https://doi.org/10.1007/s10459-015-9624-y>
- Komogortsev, O. V., & Karpov, A. (2013). Automated classification and scoring of smooth pursuit eye movements in the presence of fixations and saccades. *Behavior Research Methods*, 45, 203–215. doi:<https://doi.org/10.3758/s13428-012-0234-9>
- Krupinski, E. A. (1996). Visual scanning patterns of radiologists searching mammograms. *Academic Radiology*, 3, 137–144. doi:[https://doi.org/10.1016/S1076-6332\(05\)80381-2](https://doi.org/10.1016/S1076-6332(05)80381-2)
- Kundel, H. L., & La Follette, P. S. (1972). Visual search patterns and experience with radiological images. *Radiology*, 103, 523–528. doi:<https://doi.org/10.1148/103.3.523>
- Kundel, H. L., & Nodine, C. F. (1975). Interpreting chest radiographs without visual search. *Radiology*, 116, 527–532. doi:<https://doi.org/10.1148/116.3.527>
- Kundel, H. L., Nodine, C. F., & Carmody, D. (1978). Visual scanning, pattern recognition and decision-making in pulmonary nodule detection. *Investigative Radiology*, 13, 175–181. doi:<https://doi.org/10.1097/00004424-197805000-00001>
- Kundel, H. L., Nodine, C. F., Conant, E. F., & Weinstein, S. P. (2007). Holistic component of image perception in mammogram interpretation: Gaze-tracking study. *Radiology*, 242, 396–402. doi:<https://doi.org/10.1148/radiol.2422051997>
- Leong, J. J., Nicolaou, M., Emery, R. J., Darzi, A. W., & Yang, G. Z. (2007). Visual search behaviour in skeletal radiographs: A cross-specialty study. *Clinical Radiology*, 62, 1069–1077. doi:<https://doi.org/10.1016/j.crad.2007.05.008>
- Li, R., Shi, P. C., Pelz, J., Alm, C. O., & Haake, A. R. (2016). Modeling eye movement patterns to characterize perceptual skill in image-based diagnostic reasoning processes. *Computer Vision and Image Understanding*, 151, 138–152. doi:<https://doi.org/10.1016/j.cviu.2016.03.001>
- Loveday, T., Wiggins, M., Festa, M., Schell, D., & Twigg, D. (2013). Pattern recognition as an indicator of diagnostic expertise. In A. Fred, M. De Marsico, & G. Sanniti di Baja (Eds.), *Pattern recognition applications and methods* (pp. 1–11). Chalm, Switzerland: Springer.
- Mallett, S., Phillips, P., Fanshawe, T. R., Helbren, E., Boone, D., Gale, A., ... Halligan, S. (2014). Tracking eye gaze during interpretation of endoluminal three-dimensional CT colonography: Visual perception of experienced and inexperienced readers. *Radiology*, 273, 783–792. doi:<https://doi.org/10.1148/radiol.14132896>
- Manning, D., Ethell, S., Donovan, T., & Crawford, T. (2006). How do radiologists do it? The influence of experience and training on searching for chest nodules. *Radiography*, 12, 134–142. doi:<https://doi.org/10.1016/j.radi.2005.02.003>
- Mc Laughlin, P., Neill, S. O., Fanning, N., Mc Garrigle, A. M., Connor, O. J., Wyse, G., & Maher, M. M. (2012). Emergency CT brain: Preliminary interpretation with a tablet device: Image quality and diagnostic performance of the Apple iPad. *Emergency Radiology*, 19, 127–133.
- Milazzo, N., Farrow, D., Ruffault, A., & Fournier, J. F. (2016). Do karate fighters use situational probability information to improve decision-making performance during on-mat tasks?. *Journal of Sports Sciences*, 34, 1547–1556.
- Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, 9, 353–383.
- Nodine, C. F., Kundel, H. L., Lauver, S. C., & Toto, L. C. (1996). Nature of expertise in searching mammograms for breast masses. *Academic*

- Radiology*, 3, 1000–1006. doi:[https://doi.org/10.1016/S1076-6332\(96\)80032-8](https://doi.org/10.1016/S1076-6332(96)80032-8)
- O'Neill, E. C., Kong, Y. X., Connell, P. P., Ong, D. N., Haymes, S. A., Coote, M. A., & Crowston, J. G. (2011). Gaze behavior among experts and trainees during optic disc examination: Does how we look affect what we see? *Investigative Ophthalmology and Visual Science*, 52, 3976–3983. doi:<https://doi.org/10.1167/iovs.10-6912>
- O'Regan, J. K., & Levy-Schoen, A. (Eds.). (1987). *Eye movements: From physiology to cognition* (pp. 573–582). North Holland, Netherlands: Elsevier Science Publications BV.
- Oliva, A., & Torralba, A. (2006). Building the gist of a scene: The role of global image features in recognition. *Progress in Brain Research*, 155, 23–36.
- Over, E. A., Hooge, I. T., & Erkelens, C. J. (2006). A quantitative measure for the uniformity of fixation density: The Voronoi method. *Behavior Research Methods*, 38, 251–261.
- Rosenholtz, R. (2016). Capabilities and limitations of peripheral vision. *Annual Review of Vision Science*, 2, 437–457. doi:<https://doi.org/10.1146/annurev-vision-082114-035733>
- Sheridan, H., & Reingold, E. M. (2017). The holistic processing account of visual expertise in medical image perception: A review. *Frontiers in Psychology*, 8, 1620. doi:<https://doi.org/10.3389/fpsyg.2017.01620>
- Stockman, A. (2016). *Van variatie naar standaardisatie van detectie tijdens het radiologisch onderzoek—2011 vs. 2016* [From variation to standardization of detection during the radiological examination—2011 vs. 2016] (Master management and health care policy master thesis), Catholic University Leuven, Leuven, Belgium.
- Swenson, R. G., Hessel, S. J., & Herman, P. G. (1977). Omissions in radiology: Faulty search or stringent reporting criteria? *Radiology*, 123, 563–567. doi:<https://doi.org/10.1148/123.3.563>
- Todd, P. M., Hills, T. T., & Robbins, T. W. (2012). *Cognitive search: Evolution, algorithms, and the brain* (Vol. 9). Cambridge, MA: MIT Press.
- Treisman, A. (2006). How the deployment of attention determines what we see. *Visual Cognition*, 14, 411–443. doi:<https://doi.org/10.1080/13506280500195250>
- Van der Gijp, A., Ravesloot, C. J., Jarodzka, H., Van der Schaaf, M. F., Van der Schaaf, I. C., van Schaik, J. P., & Ten Cate, T. J. (2017). How visual search relates to visual diagnostic performance: A narrative systematic review of eye-tracking research in radiology. *Advances in Health Sciences Education*, 22, 765–787.
- van Geel, K., Kok, E. M., Dijkstra, J., Robben, S. G., & van Merriënboer, J. J. (2017). Teaching systematic viewing to final-year medical students improves systematicity but not coverage or detection of radiologic abnormalities. *Journal of the American College Radiology*, 14, 235–241. doi:<https://doi.org/10.1016/j.jacr.2016.10.001>
- Wolfe, J. M. (2012). Saved by a log: How do humans perform hybrid visual and memory search? *Psychological Science*, 23, 698–703. doi:<https://doi.org/10.1177/0956797612443968>
- Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 419–433. doi:<https://doi.org/10.1037/0096-1523.15.3.419>
- Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, 1, 0058. doi:<https://doi.org/10.1038/s41562-017-0058>
- Wolfe, J. M., Klempe, N., & Dahlen, K. (2000). Postattentive vision. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 693–716. doi:<https://doi.org/10.1037/0096-1523.26.2.693>
- Wood, G., Knapp, K. M., Rock, B., Cousens, C., Roobottom, C., & Wilson, M. R. (2013). Visual expertise in detecting and diagnosing skeletal fractures. *Skeletal Radiology*, 42, 165–172. doi:<https://doi.org/10.1007/s00256-012-1503-5>
- Yeshurun, Y., & Carrasco, M. (1999). Spatial attention improves performance in spatial resolution tasks. *Vision Research*, 39, 293–306. doi:[https://doi.org/10.1016/S0042-6989\(98\)00114-X](https://doi.org/10.1016/S0042-6989(98)00114-X)
- Ziv, G. (2017). Gaze behavior and visual attention: A review of eye tracking studies in aviation. *International Journal of Aviation Psychology*, 26, 75–104. doi:<https://doi.org/10.1080/10508414.2017.1313096>

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