From Novice to Expert: An Evaluation of the Examination Performance of Orthodontic Residents and Experienced Orthodontists of Lateral Cephalometric Images

by

Sarah E. F. Kaplan
May 16, 2018

A thesis submitted to the Faculty of the Graduate School of the University at Buffalo, State University of New York in partial fulfillment of the requirements for the degree of Masters of Orthodontics

Department of Orthodontics
# Table of Contents

List of Tables........................................................................................................V

List of Figures........................................................................................................ VI

Abstract................................................................................................................... VIII

1 Introduction..........................................................................................................1

1.1 History of Lateral Cephalometric Radiology......................................................1

1.2 Radiographic Image Interpretation.......................................................................2

1.2.1 Phases..............................................................................................................3

1.2.2 Approaches.....................................................................................................4

1.2.3 Challenges.....................................................................................................6

1.3 Current Understandings Regarding Visual Search and Expertise.........................7

1.3.1 Visual Search..................................................................................................8

1.3.2 Eye Movement Characteristics.......................................................................9

1.3.3 Time on Task.................................................................................................9

1.3.4 Visual Attention and Satisfaction of Search...................................................9

1.3.5 Teaching Strategies In Visual Search............................................................10

1.4 Eye Tracking Systems.........................................................................................12

1.5 Aim and Statement of Problem.........................................................................13

1.6 Objectives.........................................................................................................14

1.7 Hypothesis.........................................................................................................14

2 Materials and Methods.......................................................................................15

2.1 Selection of Lateral Cephalometric Images.......................................................15
3.3.2.4 Number of Revisits to AOI.............................................................38

3.3.3 Experience, AOI Visual Parameter and Image Type Interaction..........................38

3.4 Image Region Quantitative Results...........................................................................42

3.4.1 Experience and Region Visual Parameter Relationship.......................................42

3.4.2 Region Visual Parameter and Image Region Relationship....................................42

3.4.2.1 Time before First Fixation in Region.........................................................42

3.4.2.2 Number of Fixations in Region....................................................................42

3.4.2.3 Time Spent Viewing Region.......................................................................43

3.4.2.4 Number of Revisits to Region.................................................................43

3.4.3 Experience, Region Visual Parameter and Image Region Interaction....................43

3.5 Qualitative Results.................................................................................................47

3.5.1 Heat Maps Analysis..........................................................................................47

3.5.2 Scan Path Analysis...........................................................................................49

4 Discussion..................................................................................................................52

4.1 Experience Level and Overall, AOI and Region Viewing Parameter Quantitative Results Interpretation..........................................................52

4.2 Overall Viewing Parameter and Image Type Interpretation.....................................53

4.3 AOI Viewing Parameter and Abnormal Image Type Interpretation...........................54

4.4 Region Viewing Parameter and Image Region Interpretation....................................55

4.5 Interaction of Experience, Viewing Parameter and Image Type or AOI or Region Quantitative Results Interpretation..........................................................58

4.6 Evaluation of Qualitative Results............................................................................59

4.6.1 Heat Map Analysis..........................................................................................59
4.6.2 Scan Path Analysis........................................................................................................60
4.7 Study Limitations.................................................................................................................61
4.8 Implications for Orthodontic Education and Practice..........................................................64
4.9 Future Directions..................................................................................................................66
5 Conclusion.................................................................................................................................68
Appendix........................................................................................................................................70
Appendix A. SUNY University at Buffalo Institutional Review Board Approval Letter..................70
Appendix B. Google Form survey used for first found of image selection..................................71
Appendix C. Participant informed consent..................................................................................72
References......................................................................................................................................75
List of Tables

Table 1. Descriptive table of study participants. .................................................................25

Table 2. Descriptive statistics for viewing parameter by image type..................................26

Table 3. Descriptive statistics for viewing parameter in area of interest (AOI) by image type...27

Table 4. Descriptive statistics for viewing parameter by region of image.............................28

Table 5. Summary of quantitative results for parameters and image type..............................32

Table 6. Summary of quantitative results for parameters on AOI and image type....................39

Table 7. Summary of quantitative results for parameter and image region............................44
### List of Figures

**Figure 1.** Photograph of the set up (close-up) .......................................................... 19

**Figure 2.** Points of the calibration screen. ................................................................. 20

**Figure 3.** Screenshot of the Experiment Center showing the group randomization ............. 21

**Figure 4.** Photograph of the study set-up ..................................................................... 23

**Figure 5.** Results of calibration .................................................................................... 23

**Figure 6.** Mean total time (s) by image type as compared to normal ............................. 33

**Figure 7.** Mean total number of fixations by image type as compared to normal ............. 33

**Figure 8.** Mean total distance covered (cm) by image type as compared to normal ......... 34

**Figure 9.** Mean total distance covered (cm) by abnormal subcategory image type as compared to normal ................................................................. 34

**Figure 10.** Mean total number of blinks by image type as compared to normal ............... 35

**Figure 11.** Mean total number of blinks by abnormal subcategory image type as compared to normal .......................................................................................................... 35

**Figure 12.** Mean total number of saccades by image type as compared to normal ......... 36

**Figure 13.** Mean total length of saccades (s) by image type as compared to normal ......... 36

**Figure 14.** Mean time before first fixation (s) in an AOI by image type as compared to obvious .................................................................................................................. 40

**Figure 15.** Mean number of fixations in an AOI by image type as compared to obvious .... 40

**Figure 16.** Mean time spent examining (s) an AOI by image type as compared to obvious .................................................................................................................. 41

**Figure 17.** Mean number of revisits to an AOI by image type as compared to obvious ........ 41

**Figure 18.** Mean time before first fixation (s) by image region as compared to R1 .......... 45

**Figure 19.** Mean number of fixations by image region as compared to R1 ....................... 45

**Figure 20.** Mean time spent examining (s) an image region as compared to R1 ............... 46

**Figure 21.** Mean number of revisits by image region as compared to R1 ....................... 46
Figure 22. Heat map of a normal lateral cephalometric image by experience level........................................47

Figure 23. Heat map of an abnormal (obvious) lateral cephalometric image by experience level..........................48

Figure 24. Heat map of an abnormal (intermediate) lateral cephalometric image by experience level..................48

Figure 25. Heat map of an abnormal (subtle) lateral cephalometric image by experience level..........................49

Figure 26. Scan path of a normal lateral cephalometric image by experience level...........................................50

Figure 27. Scan path of an abnormal (obvious) lateral cephalometric image by experience level ......................50

Figure 28. Scan path of an abnormal (intermediate) lateral cephalometric image by experience level................51

Figure 29. Scan path of an abnormal (subtle) lateral cephalometric image by experience level........................51
Abstract

Objective: To evaluate the examination performance of orthodontic residents (OR), less experience orthodontists (LEOs, <10 years post-residency) and more experienced orthodontists (MEOs, >10 years post-residency) of lateral cephalometric images (LCIs) using an eye-tracking device.

Materials and Methods: An expert focus group selected 20 LCI’s for this observational study. A pilot study was conducted for sample size calculation for the main study. In the main study, an eye tracking system was used to accrue multiple parameters for all three groups (ORs, LEOs and MEOs), searching 5 normal and 15 LCIs of varying levels of subtlety for the presence of abnormalities. Participants included 13 ORs, 13 LEOs, and 13 MEOs. A two-way repeated measures ANOVA was performed using SPSS, and generated heat maps and scan paths were used to explore examination performance, search patterns and strategies for visual interpretation between the groups.

Results: No statistically significant difference was found for the effect of experience level (OR, LEO, MEO) on viewing parameter of an image overall, an area of interest (AOI), and image region. No statistically significant difference was found for the interaction of experience, viewing parameter, and image type. There were statistically significant differences by image type on specific overall viewing parameter (total distance covered and number of blinks), AOI viewing parameter (time before first fixation, total time examining, number of fixations and number of revisits) and region viewing parameter (time before first fixation, total time examining, number of fixations and number of revisits).

Conclusion: These findings indicate that experience does not affect overall examination performance of lateral cephalometric images, AOIs, and region, and there are no interactions affecting examination performance between the experience and image, AOI and region. Experience aside, different levels of image and AOI subtlety and lateral cephalometric image region do affect examination performance.
CHAPTER 1

1 INTRODUCTION

Lateral cephalometric images are routinely used in diagnosis and treatment planning in orthodontics. The interpretation of lateral cephalometric images is difficult because the image is a two-dimensional representation of three-dimensional anatomy. Several studies in medicine have attempted to improve understanding of the ways in which radiology specialists view images in order to improve diagnostic accuracy of novices. In dentistry, a study by Turgeon and Lam sought to explain the influence of experience and training on dental students’ examination performance regarding panoramic images. The results of these studies could be used to create methods of teaching radiologic image assessment and interpretation, using experts’ methods of analysis.

1.1 History of Lateral Cephalometric Radiology

Cephalometry has its historical roots in craniometry. The Edinburgh Encyclopedia of 1813 defined craniometry as “the art of measuring the skulls of animals so as to discover their specific differences.” The purpose of lateral cephalometric imaging is to evaluate the size of head structures, and is particularly important in the diagnosis, treatment planning, and correction of skeletal and dental abnormalities.

In early history, scientists and artist studied the cranium and other aspects of the human body. Hippocrates (460-357 BC), a pioneer in physical anthropology, first described skull forms. Centuries later, Leonardo da Vinci (1452-1519) specifically measured cephalometric features for comparisons between different skulls and heads. Albrecht Dürer published a treatise in 1528 on cranial measurements and proportions he investigated while creating artwork based on the human form. His hope was to relate anthropometry and esthetics.

Scientists of the eighteenth century sought to find relationships between head size and intelligence. Adrian van der Spigel attempted to gather true scientific measurements of the cranium. He introduced the idea of “lineae cephalometricae,” which defined four lines, the facial, occipital, frontal and sincipital lines. These lines, Spigel proposed, should be equal to one another in a well-proportioned skull. Pieter Camper (1722-1789), Deschamps (1740-1824), Daubenton (1716-1799), Sir Charles Bell (1744-1842), Johann Friedrich Blumenbach (1752-1840),
John Barclay (1758-1826), Anders Retzius (1796-1860) and other craniologists pursued their desire to learn how humans different from one another, and from animals, and why.

Advances in craniology continued into the nineteenth century. Huxley, Broca and Topinard defined new points, planes and angles, and measured different anatomical features, presented new craniometric instruments like the “craniostat” and “craniometer”, and offered new qualitative and quantitative findings of importance. 2 Around this same time, the need for standardization of methods became an important issue in craniometry. 2 In 1882, the 13th General Congress of German Anthropological Society met in Frankfurt, Germany to finalize a common method of measuring skulls. 2 It was at this meeting that the Frankfurt horizontal plane, a horizontal reference line used in several orthodontic analyses for orientation of skulls, was defined. 2,3 To top off these advances, Professor Wilhelm Conrad Roentgen discovered x-rays in 1895, which were starting to become widely used in medicine. 2

X-rays provided a means for craniologists of the twentieth century to view bones in a different way than they had before, further expanding the horizons of the field. In 1931, two dentists, Hofrath in Düsseldorf and Broadbent, separately published details on their individual apparatus to aid in the positioning and exposure of the head in relation to an x-ray source and film. 2,4 This marked the introduction of cephalometric roentgenography, cephalometric tracing and evaluation to the orthodontic specialty, and field of dentistry as a whole. 4

1.2 Radiographic Image Interpretation

Merriam-Webster defines interpretation as “the act or the result of explaining” 5 and explaining as “to make known or plain or understandable; to give reason for or cause; to show the logical development or relationships of.” 6 A combination of these two definitions leads to the very heart of what radiographic image interpretation is the understanding, and subsequent articulation, of radiographic findings.

Obtaining, tracing, and analyzing cephalometric images are routine procedures in orthodontics. Lateral cephalometric images are made for many reasons in orthodontics. These reasons include, but are not limited to, critical and descriptive expression of dimensions, accurate appraisal of jaw relationships, incisal inclination, soft tissue draping, and growth and maturation. 7,8
Although lateral cephalometric images are made to evaluate facial structures in orthodontics, their value should not be undermined as a diagnostic depiction of head and neck anatomy.  

Moffitt determined that about 50% of orthodontists would likely discover a significant, potentially life-affecting pathology on a lateral cephalometric image. It is important to ensure a complete examination and interpretation of the lateral cephalometric image in order to devise a treatment plan and accurate diagnosis of all patients.

In orthodontics, as in dentistry as a whole, specialists are typically responsible for making and interpreting their own images. Image interpretation requires one to clearly articulate their observations, and turn these observations into meanings. Radiographic interpretation of human anatomy requires a strong understanding of how an image is formed, the structure of the body, and variations of normal. Despite extensive studies over the past several decades in the field, little is known of orthodontic residents and experienced orthodontists’ approach to image interpretation.

1.2.1 Phases of Radiographic Image Interpretation

There are several phases of radiographic image interpretation. Visual search, detection of abnormal findings, feature recognition, and decision-making are steps in the analytical process of interpreting radiographic images. Blesser and Ozonoff created a model of radiographic interpretation and data processing that includes three phases: psychophysical, psychological, and nosological.

Phase one is also known as the psychophysical phase. This stage is defined by “visual interpretation.” Visual interpretation is the understanding and explaining of a patient’s two-dimensional image into that same patient’s real, three-dimensional anatomy. This system relies upon the human visual sensory system of the eye, as well as the technical aspects of the x-ray machine. An error in visual interpretation can occur as a result of poor image quality and improper viewing conditions.

Phase two is also known as the psychological phase. This stage is defined by the formation of a “visual concept.” In this phase, the brain, outside conscious awareness, processes information and organizes visual data into a meaningful image. The meaningful image created during this phase is known as the “visual concept.” Essentially, the visual concept comprises the diagnostic hypothesis of what the data means.

13
hypothesis is constantly tested by the observer, and either accepted or rejected, based upon further examination of the image.\textsuperscript{13} This process includes perception, a non-analytical process that involves rapid recognition of features.\textsuperscript{14,15} An error in perception is typically a situation in which a practitioner fails to “see” an abnormality.\textsuperscript{14,15} This is strongly affected by the expectation of the viewer, and the viewer’s training, experience, and knowledge of the case.\textsuperscript{14,15} Expert visual diagnosis is strongly shaped by the visual concept.\textsuperscript{13}

The third and final phase is known as the nosological phase. This stage is defined by a conscious decision-making effort in which the visual concept is classified into a diagnosis.\textsuperscript{13} The nosological phase uses cognitive or analytical processes, unlike the psychological phase.\textsuperscript{13,15} As one gains experience, this stage is expected to become faster and become more of a process outside conscious awareness.\textsuperscript{13} Errors in proper diagnosis may be a result of poor reasoning, misidentification, or improper concept.\textsuperscript{13}

Visual search of an image affects image interpretation. Thus, different approaches to visual search of images were developed. Tuddenham and Calvert conducted the first study to record visual search patterns of radiologists.\textsuperscript{16} Their results emphasized the importance of complete image coverage during visual search for proper image interpretation.

\textbf{1.2.2 Approaches to Radiograph Image Interpretation}

There are two primary strategies of image interpretation: feature recognition and systematic evaluation.\textsuperscript{9,14,15,17} Feature recognition is non-analytical and involves the use of “backward-reasoning.”\textsuperscript{11,18} Backward-reasoning is a process by which an abnormality is viewed as a whole, resulting a provisional diagnosis, which is then either supported or refuted through interpretation of radiographic features. Feature recognition does not require the use of conscious awareness and is predominately used by expert radiologists.\textsuperscript{11} Expert radiologists can use their experience in radiographic interpretation to make comparisons between cases, essentially recognizing patterns and perturbations from the norm. Feature recognition relies heavily upon complete visual search.\textsuperscript{16}

Significant errors in image interpretation using the feature recognition approach can occur when visual search of an image is incomplete.\textsuperscript{16} False-negative findings were due to search errors, like incomplete radiograph search, 30% of the time; Feature recognition errors, such as those with short eye fixation times and no return, accounted for 50% of errors, and decision-making errors, like long fixation times but no report of abnormality led to false-negative
findings 45% of the time. A limitation of this three-part error classification is that fixation threshold time is the defining difference between feature recognition and decision-making errors, and threshold time cannot be applied to all abnormalities on different types of images. It was hoped that systematic evaluation, the second strategy, would combat the issues associated with feature recognition.

The systematic viewing strategy aimed to ensure full image coverage in order to reduce errors in visual search using feature recognition. In contrast to feature recognition, systematic evaluation is an analytical or “forward-reasoning” approach. This method is commonly attributed to White and Pharoah. In systematic viewing, an image is first strategically divided into multiple regions. The regions of a lateral cephalogram, as described by White and Pharoah, include the base of the skull and calvarium, upper and middle face, lower face, cervical spine, airway and area of the neck, and alveolar process of the jaws and teeth. Then, the image is systematically studied, “region-by-region,” until a diagnosis can be made based upon features discovered during the search.

A study by Norman et al compared feature recognition and systematic viewing. The group trained twenty-four novice undergraduate psychology students to interpret electrocardiographs (ECG) using either forward-reasoning \( (n=16) \) or backward-reasoning \( (n=8) \). The forward-reasoning group was to identify features on the ECG data, and then generate a diagnosis, whereas, the backward-reasoning group was to generate a diagnosis, and then identify features to support their diagnosis. The study found that diagnostic accuracy was better in the backward-reasoning group (61.9%) as compared to the forward-reasoning group (49.4%). Additionally, the forward-reasoning group identified more features per ECG (5.22 features) as compared to the backward-reasoning group (3.87 features), although the backward-reasoning group identified more correct features. In all, the differences were found to be statistically significant and Normal et al concluded that forward-reasoning is detrimental when compared to backward-reasoning.

Evidence from eye-tracking studies show experienced radiologists do not adhere to a systematic search strategy and novices exhibit a focal search pattern. In fact, Tuddenham and Calvert found high intra- and inter-observer variability in radiologists’ search patterns despite the fact that the three participants were trained in the same department. Moreover, the radiologist that used systematic search missed more significant findings than the other radiologists who had used a global-focal or undefined scan pattern. This study, as well as a separate study performed by Norman et al concluded that systematic search could have effects on radiologic interpretation.
Other studies, looking to evaluate the discrepancy between instruction and actual image reading practices, have indicated that the type of search used by radiologists is influenced by the complexity and anatomical layout of the image. The global-focal model, described by Kundel and Nodine, explores image interpretation in the absence of visual search. In a “flash” experiment (200 msec), Kundel and Nodine found that even without time for visual search, radiologists detected 70% of abnormalities present in a chest image. When given unlimited time, the detection accuracy of radiologists improved to 97%. This data supported the idea that visual search begins with a general, or global, impression. Global analysis allows for gross departures from normal anatomical appearance to be detected and allows the viewer to establish content. Global analysis then guides the observer to focal analysis, in which the viewer concentrates on the abnormal areas against the norm.

Eye-movement modeling examples (EMMEs), studied by Jarodzka et al, is a fourth and recent approach to visual search and image interpretation using eye tracking. In this model, expert visual searches and scan paths are superimposed over case videos and shown to novices, guiding them towards relevant features on an image. Essentially, EMME allows novices to “see” how an expert “sees” in certain clinical situations. The results of the study showed participants studying with EMMEs had improved visual search and interpretation performance as compared to those that were not trained using EMMEs. However, no matter the model of visual search used, image interpretation is troublesome.

**1.2.3 Challenges in Radiographic Image Interpretation**

Interpretation of radiographs is complex. Significant superimposition exists when a three-dimensional structure is projected as a two-dimensional image. Interpretation is based upon complex psychophysiologic perception and cognitive processes, and is limited by human ability. Each image contains a great deal of information, which may yield potentially positive or negative findings. Visual detection, feature recognition, working memory function and cognitive reason, final interpretation, and effective communication must be used to deduce and relay a proper diagnosis.
Error in radiographic image interpretation is caused by several factors, typically inherent in the process of interpretation. Errors affecting the first phase of image interpretation may be technical factors, such as poor acquisition and sub-optimal viewing conditions. Errors affecting the second psychological or the third nosological phase may occur even with an ideal radiographic image, and may be due to lack of interpreter knowledge. Other errors may be due to complacency (attribution of a finding to the wrong cause), faulty reasoning (true-positive finding misclassified), under- or over-reading or missing findings, poor communication between radiologists and clinicians, failure to consult prior to radiograph examination, inaccurate or incomplete communication of clinical history, limited visibility of the abnormal entity, satisfaction of search (completed search early), and satisfaction of report (over-reliance on radiographic report). Thus, it is important to examine how an observer analyzes and interprets radiographic information in order to identify causes of error.

A high frequency of error and disagreement in radiologic interpretation has been reported in medical studies. A 2010 medical study which examined agreement of second readings performed by experienced abdominal imaging radiologists from Massachusetts General Hospital found that these experts disagreed with one another more than 30% of the time, and disagreed with themselves more than 25% of the time. The range of clinically significant errors reported in medical radiology, such as chest radiology, CT, barium studies, and MRI, was found to be 2-20%. A separate study found that radiologist who examined chest radiographs with lung nodules missed 27.2% of the lesions present.

Error of radiologic interpretation in dentistry has also been studied. Alsufyani and Lam found that, when comparing the diagnostic identification and interpretation of osseous dysplasia by three general dentists and three oral maxillofacial radiologists (OMRs), OMRs identified more features and interpreted 79.3% of cases correctly, as compared to general dentist who interpreted only 38.7% of cases correctly. Additionally, Rushton et al found that in panoramic radiograph interpretation, experts identified more abnormalities (20.1%) than dentists (3.5%). These studies show the reality of error and disagreement and challenges inherent in radiologic interpretation.

### 1.3 Current Understandings Regarding Visual Search and Expertise

Investigators have looked to elucidate the relationship between visual search and visual diagnostic performance. In the medical field, studies using eye-tracking technology show that expert radiologists do not see radiographs in the
same way as novices. Radiological image perception can be measured using common eye movement parameters, including first time to fixation of the abnormality, fixation durations on relevant or redundant areas, number of fixations on relevant or redundant areas, saccades, and image coverage. Visual search, time on task, eye movement characteristics, visual attention and satisfaction of search are parameters that have been used to compare novice to expert diagnostic performance.

1.3.1 Visual search

Objective measurements of eye movement are hypothesized in medical research to be direct markers of expertise. Eye movement and attention constitute important aspects of visual search.

Eye movement is complex and may reflect cognitive processes. A saccade is defined as “rapid movements of the eye from one fixation point to another.” Saccades represent eye path development. A fixation is defined as “the pause of the eye movement on a specific area of the visual field.” A fixation is typically 200-300 milliseconds. Clusters of fixations project visual information from the fovea onto the retina and to the brain. The brain is then able to process the image and provide complete clarity about what a person is looking at. It is believed that these fixation clusters represent a location of conscious attention and greater interest and engagement by the observer. Finally, blinking is described as “to close and open your eyes very quickly.”

Priority of sampling locations is made evident through examination of scan path, or the sequence by which an image is sampled. Scan paths indicate attention and can be visualized using opacity and heat maps. Gaze opacity and heat maps are eye-tracking visualizations of scan paths, or search patterns. They are qualitative displays of eye movement and visual search. These images are based on fixation counts and saccades. Eye-tracking visualizations illuminate areas where participants focused their attention and the sequence in which the image was viewed.

Gaze opacity maps, scan path sequences, and heat maps show that expert readers emphasize a global-focal radiograph interpretation method and provides evidence that experienced radiologists does not adhere to a systematic search strategy. Alternatively, qualitative evidence from eye-tracking studies shows that novices use a focal search pattern.
1.3.2 Eye Movement Characteristics

Speed of fixation on an abnormality and number of fixations are hypothesized to be direct markers of expertise. Medical studies have found that, depending on the image type (mammogram, CT, skeletal, chest or fracture X-ray) and abnormality subtlety, experts fixate on abnormalities in 0.5-5 seconds, as compared to novices who took 1.5-2 times longer to fixate on an abnormality, and up to 4.5 times longer to fixate on subtle or additional abnormalities. Conflicting results were found for saccade length and image coverage between experts and novices. Differences in eye movement characteristics are likely attributed to the approach of radiograph examination used by the two groups. Moreover, the finding is supported by the theory of “deliberate practice,” which states that repetitive practice can develop better perceptual mechanisms, yielding higher accuracy.

1.3.3 Time on Task

Visual search time, or time on task, is theorized to be indirectly related to expertise. The time it takes for an expert to view an image ranges (4-45 seconds) based on the task required, as compared to novices who take 1.5-2.5 times longer. Tasks required may include detection, interpretation or both, and may be affected by the subtlety of the lesion. Radiologists are faster and more accurate than residents in radiograph interpretation. These findings, again, support use of global-focal search strategy and “holistic recognition” process by experts. This method of interpretation allows experts to be more efficient. Myles-Worsley et al attributes radiological expertise to two aspects: knowledge of normal, and knowledge of features that signal abnormality. Few studies found no significant difference in viewing time between experts and novices.

1.3.4 Visual Attention and Satisfaction of Search

Expertise is also indirectly related to fixation duration on areas of interest when detection-only tasks are required. Fixation duration increased in time for experts when the task required both detection and interpretation. An increase in fixation duration by experts was most common in images with subtle abnormalities. Novices likely failed to recognize subtle findings as abnormal and thus did not fixate on the regions of interest for a longer time. Overall, current literature has not found statistical difference in total fixation duration between experts and novices.
Frequency of fixations is thought to be indirectly related to expertise. This is likely due to tasks required and increasing knowledge and attention to areas where abnormalities are frequently expected. Evidence from a study by Kundel and La Follette shows experienced radiologists fixate and report abnormalities correctly, while inexperienced observers fixate on abnormalities but do not recognize them.

Medical studies have also found that experts pay more attention to relevant structures (i.e. nodules on a chest x-ray) as opposed to significant anatomical structures (i.e. a heart on a chest x-ray) in radiographs. Alternatively, novices pay attention to significant structures, regardless of their relevance. Experts can identify areas of abnormality and their associated structures more readily than their novice counterparts, while rapidly passing over less important structures. These findings support the importance of visual attention in detection of abnormalities.

False negative reporting may arrive from premature termination of search, or decreased visual attention. This phenomenon is called “satisfaction of search.” Berbaum found that by adding a distractor to radiographs already containing subtle abnormalities, accuracy of detection of the native abnormalities was greatly diminished. Berbaum’s results were supported by Ashman et al, who compared detection rates in skeletal radiographs with single abnormalities to those with multiple abnormalities. Ashman found that the detection rate for second and third abnormalities in multiple abnormality cases was reduced to half. Unfortunately, false negative report and premature termination of search can have devastating consequences on diagnosis and treatment planning.

1.3.5 Teaching Strategies in Visual Search

Conflicting opinions exist in medical literature on the educational value of systematic search. Advocates of the systematic approach believe that teaching expert skills directly to learners is not effective. Rather, the process of systematic search helps novices to develop an appreciation for variations of normal anatomy. Overtime, as novices gain more experience and are able to recognize a range of normal features, the systematic search strategy is dropped. Taylor, an advocate of teaching the systematic search strategy, stated, “the impact of the training is not that readers learn the right way to move the eyes across the image but that training provides a cognitive basis for recognizing ‘perturbation’ deviations from the radiologist’s mental mode of normal appearance.”
Critics of systematic search believe that the method is not evidence-based. Several medical studies demonstrate that observers who use a systematic search strategy or checklist do not perform better than their counterparts. However, it is questionable if these findings can be related across image types, including the lateral cephalometric image.

Several medical studies have explored the effect that a systematic search strategy may have on radiologic interpretation performance. A study by Khalifa et al that sought to examine the effectiveness of training dental students and more experienced dentists in a systematic search strategy for the analysis of panoramic images. The results of the study demonstrated improved novice performance in detection and diagnosis of abnormalities in complex regions of panoramic images; however, this improvement was at the cost of over-reported findings that were typically normal variations of anatomic findings. Another study by Tuddenham and Calvert found high intra- and inter-observer variability in radiologists’ search patterns, despite their shared training background. Moreover, the radiologist that used systematic search missed more significant findings than the other radiologists, whom had used a global-focal or undefined scan pattern. This study, in addition to a study performed by Norman et al, concluded that systematic search could affect radiologic interpretation performance. Additional medical studies, which focused on chest radiographs and hand-wrist fractures, showed a discrepancy between instruction and actual radiograph reading practices. These studies indicate that the type of search used by radiologists is influenced by the complexity and anatomical layout of the image.

Complexity and anatomical layout of image play a role in the search strategy used in radiograph image interpretation. A study by Khalifa et al found that observers trained in systematic search were more likely to detect multiple abnormalities in a complex image. This particular findings highlight an important reason for understanding visual search of lateral cephalometric radiographs: Lateral cephalometric radiographs of those seeking orthodontic treatment typically contain more than one abnormal finding, ranging from obvious to subtle.

It is likely that the research findings related to hand-wrist images cannot be applied to lateral cephalometric images due to their differing levels of complexity; however, findings related to panoramic images may apply to lateral cephalometric images due to their similarities. First, lateral cephalometric images, like panoramic images, contain more anatomic features per unit area that need to be inspected than a chest and hand-wrist x-ray. A lateral cephalometric image has a high concentration of complex structures, including the cranium (frontal, parietal,
occipital, sphenoid, and ethmoid bones; anterior and posterior clinoid processes; sella turcica; floor of the anterior and posterior cranial fossa, diploe), sinuses (frontal, sphenoid, ethmoid, mastoid, maxillary), nasal bones, zygoma, maxilla, mandible, dentition, hyoid, and cervical vertebrae. Second, like panoramic images, lateral cephalometric images are complex. Superimposition of bilateral anatomical structures, effects of x-ray source and film position, and patient positioning, and more, complicate the lateral cephalometric image.\textsuperscript{3,9}

It is because of these complexities that, like panoramic images, a systematic search strategy may be successful in the detection of an abnormality, and may be advocated as a method of choice in training orthodontic residents. Several other reasons support the possible need for systematic search of lateral cephalometric images including the possibility orthodontists are more likely to fixate on the dentition and neglect the rest of the image; the chance a patient’s chief complaint may bias a dentist and draw his or her attention away from subtle abnormal findings; and the fact that not all diseases have noticeable signs and symptoms.

1.4 Eye Tracking Systems

An eye tracker is a powerful tool that allows for understanding visual attention. Eye trackers are able to perform very challenging functions and show what a person is looking at. They do not, however, have the ability to determine why someone looked. Additionally, eye trackers are able to locate an area of interest, but they are unable to determine the reason a person chose to fixate on a particular area. Finally, eye trackers are able to track a person’s gaze pattern, but not the person’s interpretation of the visual stimulus. Berbaum et al found that think-aloud data and reported viewing behavior are poor indicators of actual viewing behavior. The results of this study proved that viewing behavior is an activity that takes place outside of conscious awareness.\textsuperscript{46} Eye tracking devices allow researchers an opportunity to objectively and reliably monitor the gaze of individuals as they view radiographs, without verbal confirmation.\textsuperscript{11}

Eye tracking has been used in the past and present to understand how an individual looks at a stimulus and to subsequently learn why they looked there.\textsuperscript{11} Contemporary eye tracking devices use pupil center corneal reflection (PCCR) to track eye movements.\textsuperscript{11,50} In eye tracking devices, an infrared beam is focused on the cornea and pupil, resulting in highly visible reflections.\textsuperscript{11} A camera then captures these reflection patterns and images of the eyes,
which are synthesized and rendered by software programs to give qualitative and quantitative output. Eye trackers, and the visualizations that are generated from them, tell a great story.

Eye-tracking devices differ in both capability (sampling rates, tracking distance, accuracy and precision) and size. The capabilities of an eye-tracking device are selected based upon the purpose of use of the eye tracker and the investigation goals. In eye tracking, accuracy is described as “the distance between the actual and reported gaze positions” and precision is described as “the variance in position and latency measures.” Trackability is described as “the proportion of raw data samples that are lost during recording.” Blignaut and Wium found that accuracy, precision, and trackability of eye movement were affected significantly by ethnicity and lighting conditions.

Size differences impact the use of the eye tracker. Modern day eye trackers are miniaturized, mobile, and come with software suites capable to analyze significant amounts of data. Large devices with head stabilizers allow for accurate eye tracking data. On the other hand, smaller devices allow for portability and observer movement, but increased liberty of movement yields lower accuracy of eye tracking data.

Eye tracking can be used in dentistry to acquire both qualitative and quantitative data. The qualitative data recorded by the eye tracker includes time of analysis, gaze duration, number of fixations, saccades and blinks, and distance covered on an image. This information enables researchers to compare performance of observers. This qualitative data can then be conveyed quantitatively through visual images, such as search pattern.

The current trajectory of improvements in eye tracking software and hardware seems limitless. Eye trackers will likely become smaller in size, user-friendly, and more affordable. It is undoubtable that eye trackers will become a mainstay in clinical translational, scientific, and educational research.

1.5 **Aim and Statement of Problem**

In orthodontics, little is known of the pattern and approach to image interpretation by orthodontic residents and experienced orthodontists. Recent advances in technology have provided reliable methods of monitoring the strategy for interpreting radiographs. This study focuses on the first step of image interpretation: identification/localization of abnormality. The aim of this study is to compare the visual search and image interpretation process of lateral cephalometric images by orthodontic residents and experienced orthodontists using an eye-tracking device. Lateral
cephalometric images are complicated images and contain many important anatomical structures. Since these images are commonly used in orthodontics for diagnosis and treatment planning, it is important that residents be able to competently view, identify, and interpret the images. The findings of this study may suggest an approach of examination of lateral cephalometric images to teach to orthodontic residents so they can more quickly achieve expert status in lateral cephalometric image interpretation.

1.6 Objectives

1. To analyze lateral cephalometric assessment strategies by orthodontic residents (OR), <10 years post-residency experienced orthodontists (LEO), and ≥ 10 years post-residency experienced orthodontists (MEO) that could be applied to the education of orthodontic residents.

2. To compare groups (OR, LEO and MEO), using:

   a. Total time examining, number of fixations, distance covered, number of blinks, number and length of saccades, heat maps and scan path analysis, to find commonalities and differences in interpretation patterns of lateral cephalometric images;

   b. Time before first fixation, total time examining, number of fixations and revisits to find commonalities and differences in interpretation patterns of areas of interest on lateral cephalometric images;

   c. Time before first fixation, total time examining, number of fixations and revisits to find commonalities and differences in interpretation patterns of regions of a lateral cephalometric image.

1.7 Hypothesis

Orthodontists of varying levels of experience (OR, LEO, MEO) have similar examination performance when viewing lateral cephalometric images.
CHAPTER 2

2 MATERIALS AND METHODS

Research ethics approval for this study was obtained from SUNY University at Buffalo Research Ethics Board (STUDY00001212). The materials and methods, the study protocol, and statistical analysis procedures for this study were adapted and improved from an earlier study by Turgeon and Lam. The eye tracking system (RED-m®, Sensomotoric Instruments, Teltow, Germany) and associated computer hard-(Dell Latitude E6530, Round Rock, TX, USA) and software (Experiment Center® 3.3 Software, Sensomotoric Instruments, Teltow, Germany) were borrowed from the Faculty of Dentistry, University of Toronto, Oral and Maxillofacial Radiology Graduate Program to maintain consistency between studies.

2.1 Selection of Lateral Cephalometric Images

Over one hundred thirty-nine, digital lateral cephalometric images were pre-selected, de-identified, and decoded, from patient files at the SUNY University at Buffalo, School of Dental Medicine Department of Orthodontics in Buffalo, New York by the principal investigator. These lateral cephalometric images of patients included patients without and with head and neck pathoses. A lateral cephalometric image was rejected if it was of insufficient quality or had identifiable artifacts. An expert panel consisting of one dual-board certified orthodontist/oral and maxillofacial radiologist, one board-certified orthodontist, and one board-certified oral and maxillofacial radiologist reviewed these images. The American Board of Orthodontics and/or the American Board of Oral and Maxillofacial Radiology certified the panel members. For ease of communication and efficiency, a PDF of the de-identified pre-selected images with nine initial calibration slides was emailed to the panel for review on a personal computer. During a teleconference arranged to select the final images, the nine calibration slides were first reviewed and their difficulty level agreed-upon by the experts.

Twenty digital lateral cephalometric image of the pre-selected lateral cephalometric image group were selected for final inclusion in the study. Five of these images were normal. A lateral cephalometric image was considered normal if it was free of bony pathoses, transitional dentition, dental caries, severe periodontal disease or image artifact, and was classified as skeletal Class I (ANB = 2° ± 2°). The remaining fifteen lateral cephalometric images contained
regions of pathoses. A lateral cephalometric image was considered abnormal if it contained one or more pathoses and requires treatment. If an image was determined to be abnormal, it was then classified further into three categories: subtle, intermediate, and obvious. These categories only related to the level of difficulty to detect the abnormality, not to diagnose it. The levels of difficulty subcategories, based upon Khalifa’s classifications, were defined as: 14

**Obvious:** an abnormality that is easily identifiable in the opinion of the expert panel. All abnormalities present must be considered obvious for it to fit this category. Furthermore, it is one that everyone (including dental students, general practitioners, and specialists) in the dental community should be able to identify (i.e. severely enlarged sella turcica, pneumatization of frontal sinus)

**Intermediate:** an abnormality that may be missed by some dental students and one that a dentist or dental specialist should see with a bit of effort. At least one abnormality (regardless of if other abnormalities are more obvious) must fit the aforementioned description. This abnormality should be easily seen by an oral and maxillofacial radiologist (i.e. impacted third molar, sella turcica bridging)

**Subtle:** an abnormality that dental students and some general dentists might overlook, and one that even an oral and maxillofacial radiologist may need to pay attention to notice. At least one abnormality (regardless of if other abnormalities are present) must fit the aforementioned description (i.e. hair-on-end appearance of the skull, mucous retention pseudocyst, root resorption)

A conventional Delphi panel method was used. In this method, 100% agreement between all experts was required to accept an image. Each expert did the first round of viewing individually and separately. The results of this round were collected via Google Form survey and only visible to the primary investigator. The purpose of the survey was to simplify the agreement process for image selection. During the second round of viewing, the results of the first round were discussed. All images with 100% agreement were selected. Images in which a single expert disagreed were sent back to that expert, along with the others’ classification, and, if the expert changed his or her mind, the image was selected and distributed to the agreed-upon category of difficulty. If the expert did not agree, the image was eliminated. Images where all three experts disagreed were eliminated. After the second round of viewing, the expert panel was able to come to consensus on six normal, twenty-two abnormal (obvious), seven abnormal
(intermediate), and twelve abnormal (subtle) lateral cephalometric images. The principal investigator selected the final twenty images randomly, with five images per grouping.

AOIs identified by the focus group as being present in the images included long condylar neck, mandibular prognathism, severe maxillary/mandibular anterior proclination, taurodontism, sella turcica bridging, erupting/impacted teeth (canine, second molar, third molar), elongated midface, root resorption, fused cervical vertebrae, impacted second molar, excessive overjet, excessive overbite, pneumatized frontal sinus, mucous retention pseudocyst, frontal bossing, enlarged sella turcica, severe high/low mandibular plane angle, maxillary hypoplasia, negative overjet, enlarged palatine tonsil, hair-on-end appearance of skull, reduced diploic space, and supernumerary teeth. Although these abnormalities exist within a range of normal, only those that were “outside” that range of normal, and may require orthodontic/surgical treatment, were selected and also categorized based upon their ability to be identified by a given experience level.

2.2 Sample Size

A sample size calculation needed to be performed prior to the beginning of the study because only one similar study, by Turgeon and Lam, of this nature exists. Their study had based its sample size off of earlier publications in the medical field and conducted a pilot study with ten novice participants for a sample size calculation. A similar approach was taken in this investigation. A pilot study was performed with 5 novice participants and 4 expert participants. Calculations were completed using the total time examining the image [one per category (normal, obvious, intermediate and subtle)] and their associated mean and standard deviations for novices and experts. Total time examining an image was the selected parameter that defined novices from experts based up on findings from earlier studies. Sample size estimation showed the need for a minimum of 11 subjects per group to attain statistical significance between observer groups ($\alpha = 0.05, 1-\beta = 0.9$) if the standard deviation for total time examining an image was 20 seconds.

2.3 Observer Selection

Orthodontic residents from the first (estimated graduation year: 2020), second (estimated graduation year: 2019), and third year (estimated graduation year: 2018) of the SUNY University at Buffalo Department of Orthodontics
were recruited. This group was selected because they had received training on lateral cephalometric interpretation through the SUNY University at Buffalo Department of Orthodontics. They have also had experience in viewing and interpreting lateral cephalometric images. A small portion of these resident groups was foreign-trained dentists or had private practice experience; however, they were considered as orthodontic residents. Differentiation was made during the test between these residents and the regular stream orthodontic residents in case comparison between groups was to be made.

The expert group included orthodontists, without and with board certification, of varying years out of residency. Currently 42% of the American Association of Orthodontists (AAO) are board certified, and 88% of orthodontists in the United States are members of the AAO.

The recruitment for pilot study participants began one week prior to the pilot study commencement by sending out an email to UB SDM orthodontic faculty and residents. The recruitment for main study participant began one week prior to the main study commencement by sending out an email to the Greater Buffalo Association of Orthodontists (GBAO), as well as orthodontic residents and faculty that had not participated in the pilot study. Some participants responded immediately with their intention to participate. Further recruitment of subjects was performed using word of mouth.

2.4 Eye Tracker System

The RED-m® (Sensomotoric Instruments, Teltow, Germany) eye tracker system was used to maintain consistency with an earlier published study. This small eye tracker (24 cm by 2.5 cm by 3.3 cm, weighing 130 g) offers a sampling rate of 120 Hz. A sampling rate of 120 Hz allows for collection of 120 data points for each of the tracked eyes. This rate is sufficient for subjects to maintain a steady head position during the study and does not require the use of a head stabilization device. This was ideal for our study because it minimized the potential for subjects to think about the eye tracker during the study. The operating distance between the device and an observer’s eyes was between 50 cm and 75 cm. The optimal operating distance is between 60 cm to 65 cm because the tracking area of the system at this distance is 32 cm by 21 cm. This tracking area allows for some head movement (maximum head movement velocity = 15cm/s at 60Hz). Furthermore, the system has a gaze position accuracy of 0.5° and a spatial resolution of 0.1°. At 65 cm, a 1° change corresponds to 11 mm. Thus, the accuracy of this
system represents an error of approximately 5 mm. Finally, the system tracks both eyes and works with most glasses and lenses.

The eye tracker was mounted using a magnetic strip onto the base of a 15.6-inch laptop screen (Dell Latitude E6530, Round Rock, TX, USA) with a display resolution of 1600 by 900 pixels. The laptop was used to collect participant responses to questions regarding each image. The principle investigator used a separate screen (19-inch Dell 1909 Wf, Round Rock, TX, USA; Display resolution of 1440 by 900 pixels), connected to the laptop via VGA cable, to ensure that subjects stayed within the tracking range and operating distance of the eye tracker. (Figure 1)

Figure 1. Photograph of the setup (close-up).

Note: Photograph taken by Sarah Kaplan.

Experiment Center ® 3.3 software is associated with the Red-m eye tracker system (Sensomotoric Instruments, Teltow, Germany) and is used to build experiments with images, amongst other types of stimuli. The software allows for order randomization of the twenty included images and associated questions for each participant. A nine-point initial on-screen calibration was used for each participant in the study, followed by a four-point confirmation calibration. (Figure 2)
Figure 2. Points of the calibration screen.

Note: Photograph taken by Sarah Kaplan.

2.4.1 Software Programming

The principal investigator (S.K.) customized the program software for the purposes of the study. (Figure 3) For each study, the first seven screens were used to obtain epidemiologic information about each participant (sex, age, educational background, board certification, years practicing as an orthodontist, constituent orthodontic organization, and school of orthodontic training). After these slides, the next one hundred screens consisted of the twenty final lateral cephalometric images and their associated questions, randomized as a group. The associated questions were:

i. Is there an abnormality on this image?

ii. If there is an abnormality, where is it or are they located?

iii. If there is an abnormality, what is your interpretation/diagnosis?

iv. On a scale from 1 to 10 (1 = ‘not confident’; 10 = ‘completely confident’), rate your confidence level
with regards to your interpretation/diagnosis of this image.

The last screen thanked participants and let them know the study was finished.

**Figure 3.** Screenshot of the Experiment Center showing the group randomization.

*Note:* Photograph taken by Sarah Kaplan.
2.5 Study Protocol

Both the pilot and main studies were primarily conducted in a well-lit venue at SUNY University at Buffalo, School of Dental Medicine in Buffalo, New York. The space was lit with fluorescent lights that did not allow dimming. In situations where the participant was unable to travel to SUNY UB SDM for the main study, the study equipment was brought to the select off-campus office location and a room with similar requirement was used to perform the study. Prior to positioning the participant, the protocol of the study was explained. Participants were asked to view images as if they were in their office preparing a preliminary diagnosis prior to a consultation. If the subject agreed to participation, a consent form was signed, followed by an explanation of the eye tracker, tips on how to ensure optimal gaze tracking (being centered within the field of the camera at a distance of approx. 60-65 cm) and how to operate the software.

The participant was seated with the laptop screen placed approximately 60 to 65 cm. The operator or principal investigator was seated at 90° to the subject. (Figure 4) The purpose of this seating arrangement was to monitor the subject’s gaze, head position, and movement on the second monitor.

Each participant was identified in the eye tracking software with a code, “TS” for novices and “TF” for experts, followed by two-digit number. After an initial calibration, the software showed the gaze tracking accuracy for the participant, and the operator could accept the calibration or reject it and elect to recalibrate the system. An accuracy of less than one degree was considered accurate. (Figure 5)

The operator gave feedback to the subject about missed abnormalities or interpretation, recurrent mistakes, and answered any other questions the subject had once the study was completed. The participants were asked to not discuss the cases with other potential participants.

The data for each participant was exported from BeGaze ® (Sensomotoric Instruments, Teltow Germany) to Excel ® (Microsoft Corp., Redomond, WA, USA) where they were grouped according to participant type [novice (OR) or expert (EO)], whether the radiograph was normal or abnormal, and level of difficulty (obvious, intermediate, subtle). The participants were then sub-grouped into orthodontic residents (novices, OR), less experienced orthodontists
(orthodontists with <10 years of post-residency experience, LEO) and more experienced orthodontists (orthodontists with $\geq10$ years of post-residency experience, MEO).

**Figure 4.** Photograph of study setup.

![Image of study setup](image1.png)

*Note:* The observer is to the left and the participant is to the right. Photograph taken by Sarah Kaplan.

**Figure 5.** Results of calibration.

![Image of calibration results](image2.png)

*Note:* Photograph taken by Sarah Kaplan.
2.6 Statistical Analysis

The primary outcome measures of the study were the viewing parameters and included total viewing time (seconds, s), number of fixations, distance covered (centimeters, cm), number of blinks, number of saccades, length of saccades (milliseconds, ms), time before first fixation in an area of interest (AOI) (seconds, s), number of fixations in AOI, total time spent in AOI (seconds, s), number of revisits to an AOI, time before first fixation in a region (seconds, s), number of fixations in a region, total time spent in a region (seconds, s), and number of revisits to a region. These measures were selected because they indicate how a participant cognitive function of visual search: abnormality localization and interpretation.

Data were analyzed using IBM SPSS Statistics for Windows, Version 24. Descriptive statistics were calculated for all variables. A two-way repeated measures analysis of variance (ANOVA) was used to compare the three experience groups (OR, LEO, and MEO), the viewing parameter, and the image type (normal, abnormal, obvious, intermediate, subtle) or region of image [Alveolar bone and teeth (R1), base of skull and calvarium (R2), upper and middle face (R3), lower face (R4) and cervical spine, airway and neck (R5)] The data were analyzed at the 5% significance levels. If there was no significant differences for the evaluated parameter between normal and abnormal images types, no evaluation of the abnormal sub-classified images (obvious, intermediate or subtle) was performed; however, if a statistically significant difference was found, the abnormal image sub-classification (obvious, intermediate, subtle) was further investigated to see which image type differed from the normal image using pairwise comparisons.

Qualitative findings were presented and studied using heat maps and gaze plots generated by the eye tracking software. Heat maps show “how looking is distributed over the stimulus.” Gaze plots show the “location, order and time spent looking at locations on the stimulus.” Heat maps for both groups for every radiograph were saved as Joint Photographic Experts Group (JPEG) files. Scan paths were also saved as JPEG files for each participant for every radiograph.
CHAPTER 3

3 RESULTS

3.1 DESCRIPTIVE STATISTICS (Table 1-4)

Descriptive statistics were collected for each group (OR, LEO, MEO). A convenient sample of people willing to participate was studied and amounted to a total of 13 OR, 13 LEO, and 13 MEO. (Tables 1-4)

In two cases, there was no data reported by the program, so the data for those particular images for those particular participants is based on one less instance.

Table 1. Descriptive table of study participants.

<table>
<thead>
<tr>
<th>Group</th>
<th>Gender</th>
<th>Level of Education</th>
<th>ABO Certified</th>
<th>Residency Constituent Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orthodontic Residents</td>
<td>M/F</td>
<td>Certificate of Orthodontics</td>
<td>Dual-Certified (Masters/Certificate) and Other</td>
<td>Yes/No</td>
</tr>
<tr>
<td></td>
<td>6/7</td>
<td>Not Applicable</td>
<td>Not Applicable</td>
<td>13 NESO</td>
</tr>
<tr>
<td>Experienced Orthodontists</td>
<td>&lt;10 Years</td>
<td>9/4</td>
<td>3</td>
<td>9 (Dual-Certified) 1 (Craniofacial Fellowship)</td>
</tr>
<tr>
<td></td>
<td>≥10 years</td>
<td>12/1</td>
<td>8</td>
<td>5 (Dual-Certified)</td>
</tr>
</tbody>
</table>

Table 2. Descriptive statistics for viewing parameter by image type.

<table>
<thead>
<tr>
<th>Viewing Parameter</th>
<th>Image Type</th>
<th>Participants</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Orthodontic Residents</td>
<td>Experienced Orthodontists</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Under 10</td>
<td>10 and Over</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Time (s)</td>
<td>Normal</td>
<td>55.36</td>
<td>52.50</td>
<td>46.21</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abnormal</td>
<td>58.10</td>
<td>53.22</td>
<td>50.37</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Obvious</td>
<td>62.40</td>
<td>57.59</td>
<td>54.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>54.02</td>
<td>49.98</td>
<td>48.42</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Subtle</td>
<td>57.88</td>
<td>52.07</td>
<td>48.68</td>
<td></td>
</tr>
<tr>
<td>Number of Fixations</td>
<td>Normal</td>
<td>138</td>
<td>125</td>
<td>124</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abnormal</td>
<td>142</td>
<td>130</td>
<td>132</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Obvious</td>
<td>153</td>
<td>142</td>
<td>139</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>131</td>
<td>119</td>
<td>128</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Subtle</td>
<td>142</td>
<td>128</td>
<td>129</td>
<td></td>
</tr>
<tr>
<td>Distance Covered (cm)</td>
<td>Normal</td>
<td>443.72</td>
<td>387.70</td>
<td>364.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abnormal</td>
<td>448.58</td>
<td>412.21</td>
<td>399.83</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Obvious</td>
<td>483.21</td>
<td>445.19</td>
<td>417.62</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>421.19</td>
<td>377.45</td>
<td>391.82</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Subtle</td>
<td>441.33</td>
<td>413.98</td>
<td>390.05</td>
<td></td>
</tr>
<tr>
<td>Number of Blinks</td>
<td>Normal</td>
<td>20</td>
<td>18</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abnormal</td>
<td>23</td>
<td>19</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Obvious</td>
<td>25</td>
<td>21</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>21</td>
<td>19</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Subtle</td>
<td>23</td>
<td>17</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Number of Saccades</td>
<td>Normal</td>
<td>147</td>
<td>136</td>
<td>127</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abnormal</td>
<td>154</td>
<td>142</td>
<td>135</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Obvious</td>
<td>164</td>
<td>154</td>
<td>144</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>142</td>
<td>133</td>
<td>130</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Subtle</td>
<td>154</td>
<td>139</td>
<td>131</td>
<td></td>
</tr>
<tr>
<td>Length of Saccades (ms)</td>
<td>Normal</td>
<td>26.72</td>
<td>29.55</td>
<td>34.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abnormal</td>
<td>27.71</td>
<td>29.25</td>
<td>36.43</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Obvious</td>
<td>27.08</td>
<td>29.95</td>
<td>41.49</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>27.61</td>
<td>30.46</td>
<td>34.35</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Subtle</td>
<td>28.45</td>
<td>27.35</td>
<td>33.45</td>
<td></td>
</tr>
</tbody>
</table>

*Note:* Values shown represent means. “Abnormal” represents a mean of “obvious,” “intermediate,” and “subtle” data.
Table 3. Descriptive statistics for viewing parameter in area of interest (AOI) by image type.

<table>
<thead>
<tr>
<th>Viewing Parameter</th>
<th>Image Type</th>
<th>Participants</th>
<th>Orthodontic Residents</th>
<th>Experienced Orthodontists</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Under 10</td>
<td>10 and Over</td>
</tr>
<tr>
<td>Time before 1st Fixation (s)</td>
<td>Abnormal</td>
<td>12.92</td>
<td>12.47</td>
<td>12.39</td>
</tr>
<tr>
<td></td>
<td>Obvious</td>
<td>10.97</td>
<td>10.89</td>
<td>9.17</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>11.99</td>
<td>10.21</td>
<td>12.07</td>
</tr>
<tr>
<td></td>
<td>Subtle</td>
<td>15.81</td>
<td>16.29</td>
<td>15.94</td>
</tr>
<tr>
<td>Number of Fixations in AOI</td>
<td>Abnormal</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Obvious</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>6</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Subtle</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Total Time in AOI (s)</td>
<td>Abnormal</td>
<td>2.54</td>
<td>2.92</td>
<td>2.48</td>
</tr>
<tr>
<td></td>
<td>Obvious</td>
<td>3.62</td>
<td>3.80</td>
<td>3.90</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>2.25</td>
<td>2.59</td>
<td>1.90</td>
</tr>
<tr>
<td></td>
<td>Subtle</td>
<td>1.76</td>
<td>2.38</td>
<td>1.63</td>
</tr>
<tr>
<td>Number of Revisits</td>
<td>Abnormal</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Obvious</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Subtle</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

**Note:** Values shown represent means. “Abnormal” represents a mean of “obvious,” “intermediate,” and “subtle” data.
Table 4. Descriptive statistics for viewing parameter by region of image.

<table>
<thead>
<tr>
<th>Viewing Parameter</th>
<th>Radiograph Region</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Orthodontic Residents</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Under 10</td>
</tr>
<tr>
<td>Time before 1st Fixation (s)</td>
<td>Alveolar Bone and Teeth</td>
<td>2.345</td>
</tr>
<tr>
<td></td>
<td>Base of Skull and Calvarium</td>
<td>0.076</td>
</tr>
<tr>
<td></td>
<td>Upper and Middle Face</td>
<td>1.793</td>
</tr>
<tr>
<td></td>
<td>Lower Face</td>
<td>3.623</td>
</tr>
<tr>
<td>Number of Fixations in AOI</td>
<td>Alveolar Bone and Teeth</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>Base of Skull and Calvarium</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Upper and Middle Face</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>Lower Face</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Airway, Cervical Spine, Neck</td>
<td>19</td>
</tr>
<tr>
<td>Total Time in AOI (s)</td>
<td>Alveolar Bone and Teeth</td>
<td>14.956</td>
</tr>
<tr>
<td></td>
<td>Base of Skull and Calvarium</td>
<td>8.825</td>
</tr>
<tr>
<td></td>
<td>Upper and Middle Face</td>
<td>13.887</td>
</tr>
<tr>
<td></td>
<td>Lower Face</td>
<td>10.187</td>
</tr>
<tr>
<td>Number of Revisits</td>
<td>Alveolar Bone and Teeth</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Base of Skull and Calvarium</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Upper and Middle Face</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Lower Face</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Airway, Cervical Spine, Neck</td>
<td>5</td>
</tr>
</tbody>
</table>

*Note:* Values shown represent means. “Abnormal” represents a mean of “obvious,” “intermediate,” and “subtle” data.
3.2 OVERALL IMAGE QUANTITATIVE RESULTS (Table 5, Figures 6-13)

3.2.1 Experience and Visual Parameter Relationship

Regardless of lateral cephalometric image type [normal (n) or abnormal (a)], there was no statistically significant difference between experience level (e, ie: OR, LEO, MEO) and total time to examine ($p_e(TotalTime) = 0.724$), total number of fixations ($p_e(Fixations) = 0.879$), total distance covered ($p_e(Distance) = 0.710$), total number of blinks ($p_e(Blinks) = 0.146$), total number of saccades taken ($p_e(Saccades) = 0.807$), and total average length of saccades ($p_e(SaccadeLength) = 0.586$). Thus, the null hypotheses cannot be rejected and the participants could be grouped into a single group of participants, regardless of experience level. (Table 5)

3.2.2 Visual Parameter and Image Type Relationship

3.2.2.1 Total Time, Total Number of Fixations, Total Number of Saccades, Total Length of Saccades

Experience level aside, there was no statistically significant difference between image type and total time to examine ($p_n(TotalTime) = 0.084$), total number of fixations ($p_n(Fixations) = 0.093$), total number of saccades ($p_n(Saccades) = 0.066$), and total length of saccades ($p_n(SaccadeLength) = 0.248$). Thus, the null hypotheses cannot be rejected and no further exploration of abnormal image sub-classification was needed. (Table 5, Figures 6-7, Figures 12-13)

3.2.2.2 Total Distance Covered

There was statistically significant difference between total distance covered on a lateral cephalometric image and image type, regardless of experience level, $F(1, 36) = 5.872, p_{na}(Distance) = 0.021$. For normal images there was an average of 398.472 cm covered and for abnormal images there was an average of 420.205 cm covered. Because a statistically significant difference between n and a images existed, the abnormal image sub-classification [obvious (o), intermediate (i), or subtle (s)] was investigated to find where the difference was. (Table 5, Figure 8)

Similar to examination between n and a images, regardless of abnormal image sub-classification (o, i, s), no statistically significant difference between experience level and total distance covered on a lateral cephalometric
image existed, $p_e(Distance) = 0.765$. Thus the null hypotheses cannot be rejected and the participants could stay grouped as a single group of participants, regardless of experience level. (Table 5)

Furthermore, there was no statistically significant difference between total distance covered on a lateral cephalometric image and image type, regardless of experience level, $p_{nois}(Distance) = 0.057$. Thus the null hypothesis cannot be rejected. Although this exploratory analysis is not significant, the finding does show a trend with increased total distance for obvious images, followed by subtle and intermediate. For obvious images there was an average of 448.673 cm covered, for intermediate images there was an average of 396.822 cm covered, and for subtle images there was an average of 415.120 cm covered. (Table 5, Figure 9)

### 3.2.2.3 Total Number of Blinks

There was a statistically significant difference between total number of blinks on a lateral cephalometric image and image type, regardless of experience level, $F (1, 36) = 8.506, p_{no}(Blinks) = 0.006$. For normal images there was an average of 16 blinks and for abnormal images there was an average of 18 blinks. Thus the null hypotheses can be rejected. Because a statistically significant difference between $n$ and $a$ images existed, the abnormal image sub-classification was investigated to find where the difference was. (Table 5, Figure 10)

Similar to examination between $n$ and $a$ images, regardless of abnormal image sub-classification ($o$, $i$, $s$), no statistically significant difference between experience level and total number of blinks on a lateral cephalometric image existed, $p_e(Blinks) = 0.178$. Thus the null hypotheses cannot be rejected and the participants could stay grouped as a single group of participants, regardless of experience level, for abnormal image sub-classification. (Table 5)

There was a statistically significant difference between total number of blinks on a lateral cephalometric image and image type, regardless of experience level, $F (3, 34) = 3.045, p_{nois}(Blinks) = 0.042$. For obvious images there was an average of 20 blinks, for intermediate images there was an average of 17 blinks and for subtle images there was an average of 17 blinks. A statistically significant difference in total number of blinks was found between normal and obvious images ($p_{no}(Blinks) = 0.005$) but not between normal and intermediate ($p_{ni}(Blinks) = 0.298$) and normal and subtle images ($p_{ns}(Blinks) = 0.178$). Thus the null hypotheses can be rejected. (Table 5, Figure 11)
3.2.3 Experience, Visual Parameter and Image Type Interaction

Interactions were examined because of the possible combined effects of experience and image type on the visual parameter. There was no statistically significant interaction between experience level, image type, and visual parameter in all instances examined ($p_{e, na} (\text{TotalTime}) = 0.618$, $p_{e, na} (\text{Fixations}) = 0.851$, $p_{e, na} (\text{Distance}) = 0.372$, $p_{e, nois} (\text{TotalTime}) = 0.722$, $p_{e, na} (\text{Blinks}) = 0.511$, $p_{e, nois} (\text{Blinks}) = 0.805$, $p_{e, na} (\text{Saccades}) = 0.968$, $p_{e, na} (\text{SaccadeLength}) = 0.468$). Thus the null hypotheses cannot be rejected. (Table 5)
Table 5. Summary of quantitative results for parameters and image type.

<table>
<thead>
<tr>
<th>Viewing Parameter</th>
<th>Image Type</th>
<th>Participants</th>
<th>$p$-value</th>
<th>$e$ (vp)</th>
<th>$i$ (vp)</th>
<th>$e, i$ (vp)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total Time (s)</strong></td>
<td>Normal</td>
<td>51.356</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abnormal</td>
<td>53.895</td>
<td>0.724</td>
<td>0.084</td>
<td>0.618</td>
<td></td>
</tr>
<tr>
<td><strong>Number of Fixations</strong></td>
<td>Normal</td>
<td>129</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abnormal</td>
<td>135</td>
<td>0.879</td>
<td>0.093</td>
<td>0.851</td>
<td></td>
</tr>
<tr>
<td><strong>Distance Covered (cm)</strong></td>
<td>Normal</td>
<td>398.472</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abnormal</td>
<td>420.205</td>
<td>0.710</td>
<td>0.021</td>
<td>0.372</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abnormal Image Sub-Classification Results</td>
<td></td>
<td></td>
<td>0.765</td>
<td>0.057</td>
<td>0.722</td>
</tr>
<tr>
<td></td>
<td>Obvious</td>
<td>448.673</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>396.822</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Subtle</td>
<td>415.120</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Number of Blinks</strong></td>
<td>Normal</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abnormal</td>
<td>18</td>
<td>0.146</td>
<td><strong>0.006</strong></td>
<td>0.511</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abnormal Image Sub-Classification Results</td>
<td></td>
<td></td>
<td><strong>0.178</strong></td>
<td><strong>0.042</strong></td>
<td>0.805</td>
</tr>
<tr>
<td></td>
<td>Obvious</td>
<td>20</td>
<td></td>
<td><strong>0.005</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>17</td>
<td></td>
<td>0.298</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Subtle</td>
<td>17</td>
<td></td>
<td>0.178</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Number of Saccades</strong></td>
<td>Normal</td>
<td>137</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abnormal</td>
<td>143</td>
<td>0.807</td>
<td>0.066</td>
<td>0.968</td>
<td></td>
</tr>
<tr>
<td><strong>Length of Saccades (ms)</strong></td>
<td>Normal</td>
<td>30.111</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abnormal</td>
<td>31.133</td>
<td>0.586</td>
<td>0.248</td>
<td>0.468</td>
<td></td>
</tr>
</tbody>
</table>

*Note:* All image types were compared to normal images. Values shown represent means. Statistically significant differences ($p<0.05$) are in red. $e =$ experience, $vp =$ visual parameter, $i =$ image type.
**Figure 6.** Mean total time examining (s) by image type as compared to normal.

![Total Time Examining by Image Type](chart1)

*Note:* These image types are not statistically significantly different.

**Figure 7.** Mean total number of fixations by image type as compared to normal.

![Total Number of Fixations by Image Type](chart2)

*Note:* These image types are not statistically significantly different.
Figure 8. Mean total distance covered (cm) by image type as compared to normal.

![Bar chart showing total distance covered by image type]

*Note: “*” denotes a statistically significant difference ($p<0.05$). Bars connect image types that are statistically significantly different.

Figure 9. Mean total distance covered (cm) by abnormal subcategory image type as compared to normal.

![Bar chart showing total distance covered by abnormal subcategory image type]

*Note: These image types are not statistically significantly different.
Figure 10. Mean total number of blinks by image type as compared to normal.

![Total Number of Blinks by Image Type](image10.png)

*Note:* “*” denotes a statistically significant difference (*p* < 0.05). Bars connect image types that are statistically significantly different.

Figure 11. Mean total number of blinks by abnormal subcategory image type as compared to normal.

![Total Number of Blinks by Image Type](image11.png)

*Note:* “*” denotes a statistically significant difference (*p* < 0.05). Bars connect image types that are statistically significantly different.
**Figure 12.** Mean total number of saccades by image type as compared to normal.

![Total Number of Saccades by Image Type](image)

*Note:* These image types are not statistically significantly different.

**Figure 13.** Mean total length of saccades (s) by image type as compared to normal.

![Total Duration of Saccades by Image Type](image)

*Note:* These image types are not statistically significantly different.
3.3  ABNORMAL IMAGE AREA OF INTEREST (AOI) QUANTITATIVE RESULTS

(Table 6, Figures 14-17)

3.3.1 Experience and Visual Parameter Relationship

Regardless of abnormal sub-classification image type \((o, i, s)\), there was no statistically significant difference between experience level and time before first fixation on an AOI \((p_{e(AOIEntry)} = 0.968)\), number of fixations in an AOI \((p_{e(AOIFixations)} = 0.992)\), time spent viewing an AOI \((p_{e(AOITimeSpent)} = 0.604)\), and number of revisits to an AOI \((p_{e(AOIRevisits)} = 0.973)\). Thus the null hypotheses cannot be rejected and the participants could be grouped into a single group of participants, regardless of experience level. (Table 6)

3.3.2 AOI Visual Parameter and Image Type Relationship

3.3.2.1 Time Before First Fixation in AOI

Experience level aside, there was a statistically significant difference between time before first fixation on an AOI on a lateral cephalometric image and abnormal sub-classification image type, regardless of experience level, \(F(2, 35) = 6.976, p_{ois(AOIEntry)} = 0.003\). For obvious images, there was an average of 10.346 seconds before first fixation in an AOI. For intermediate images, there was an average of 11.423 seconds before first fixation in an AOI. For subtle images, there was an average of 15.808 seconds before first fixation in AOI. A statistically significant difference in time before first fixation was found between obvious and subtle images \((p_{ois(AOIEntry)} = 0.002)\) but not between obvious and intermediate images \((p_{ois(AOIEntry)} = 0.403)\). Thus the null hypothesis can be rejected. (Table 6, Figure 14)

3.3.2.2 Number of Fixations in AOI

There was a statistically significant difference between number of fixations in an AOI on a lateral cephalometric image and image type, regardless of experience level, \(F(2, 35) = 13.937, p_{ois(AOIFixations)} < 0.001\). For obvious images, there was an average of 10 fixations in an AOI. For intermediate images, there was an average of 6 fixations in an AOI. For subtle images, there was an average of 5 fixations \((SE = 1)\) in an AOI. A statistically significant difference
in number of fixations in AOI was found between obvious and intermediate ($p_{oi} (AOIFixations) < 0.001$), and obvious and subtle images ($p_{oi} (AOIFixations) < 0.001$). Thus the null hypothesis can be rejected. (*Table 6, Figure 15*)

### 3.3.2.3 Time Spent Viewing AOI

There was a statistically significant difference between time spent viewing an AOI on a lateral cephalometric image and image type, regardless of experience level, $F (2, 35) = 20.523, p_{ois} (AOITimeSpent) < 0.001$. For obvious images, there was an average of 3.774 seconds spent in an AOI. For intermediate images, there was an average of 2.245 seconds spent in an AOI. For subtle images, there was an average of 1.921 seconds spent in an AOI. A statistically significant difference in time spent viewing an AOI was found between obvious and intermediate ($p_{oi} (AOITimeSpent) < 0.001$), and obvious and subtle images ($p_{ois} (AOITimeSpent) < 0.001$). Thus the null hypotheses can be rejected. (*Table 6, Figure 16*)

### 3.3.2.4 Number of Revisits to AOI

There was a statistically significant difference between number of revisits to an AOI on a lateral cephalometric image and image type, regardless of experience level, $F (2, 35) = 18.081, p_{ois} (AOIRevisits) < 0.001$. For obvious images, there was an average of 5 revisits to an AOI. For intermediate images, there was an average of 4 revisits to an AOI. For subtle images, there was an average of 2 revisits to an AOI. A statistically significant difference in number of revisits to an AOI was found between obvious and intermediate ($p_{oi} (AOIRevisits) = 0.011$), and obvious and subtle images ($p_{ois} (AOIRevisits) < 0.001$). Thus the null hypotheses can be rejected. (*Table 6, Figure 17*)

### 3.3.3 Experience, AOI Visual Parameter and Image Type Interaction

There was no statistically significant interaction between experience level, image type, and visual parameter in all instances examined ($p_{e, ois} (AOIEntry) = 0.773, p_{e, ois} (AOIFixations) = 0.510, p_{e, ois} (AOITimeSpent) = 0.759$, and $p_{e, ois} (AOIRevisits) = 0.885$). Thus the null hypotheses cannot be rejected. (*Table 6*)
Table 6. Summary of quantitative results for parameters on AOI and image type.

<table>
<thead>
<tr>
<th>Viewing Parameter</th>
<th>Image Type</th>
<th>Participants</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$e$ ($vp$)</td>
</tr>
<tr>
<td>Time before 1st Fixation (s)</td>
<td>Obvious</td>
<td>10.346</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abnormal Image Sub-Classification</td>
<td>0.968</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>11.423</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Subtle</td>
<td>16.012</td>
<td></td>
</tr>
<tr>
<td>Number of Fixations in AOI</td>
<td>Obvious</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abnormal Image Sub-Classification</td>
<td>0.992</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Subtle</td>
<td>5</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Total Time in AOI (s)</td>
<td>Obvious</td>
<td>3.774</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abnormal Image Sub-Classification</td>
<td>0.604</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>2.245</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Subtle</td>
<td>1.921</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Number of Revisits</td>
<td>Obvious</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abnormal Image Sub-Classification</td>
<td>0.973</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>4</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>Subtle</td>
<td>2</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Note: All image types were compared to obvious images. Values shown represent means. Statistically significant differences ($p<0.05$) are in red. $e$ = experience, $vp$ = visual parameter, $i$ = image type.
**Figure 14.** Mean time before first fixation (s) in an AOI by image type as compared to obvious.

![Time Before First Fixation in AOI](chart1)

**Note:** “*” denotes a statistically significant difference ($p<0.05$). Bars connect image types that are statistically significantly different.

**Figure 15.** Mean number of fixations in an AOI by image type as compared to obvious.

![Number of Fixations in AOI](chart2)

**Note:** “*” denotes a statistically significant difference ($p<0.05$). Bars connect image types that are statistically significantly different.
Figure 16. Mean time spent examining (s) an AOI by image type as compared to obvious.

![Time Spent Examining AOI](image16)

*Note:* “ * ” denotes a statistically significant difference ($p<0.05$). Bars connect image types that are statistically significantly different.

Figure 17. Mean number of revisits to an AOI by image type as compared to obvious.

![Number of Revisits to AOI](image17)

*Note:* “ * ” denotes a statistically significant difference ($p<0.05$). Bars connect image types that are statistically significantly different.
3.4 IMAGE REGION QUANTITATIVE RESULTS (*Table 7*, *Figures 18-21*)

3.4.1 Experience and Region Visual Parameter Relationship

Regardless of image region [alveolar bone and teeth (*R1*), base of skull and calvarium (*R2*), upper and middle face (*R3*), lower face (*R4*), cervical spine, airway, area of the neck (*R5*)], there was no statistically significant difference between experience level and time before first fixation in a region (*p* (*RegionEntry*) = 0.501), number of fixations in a region (*p* (*RegionFixation*) = 0.360), time spent viewing a region (*p* (*RegionTimeSpent*) = 0.163), and number of revisits to a region (*p* (*RegionRevisits*) = 0.418). Thus the null hypotheses cannot be rejected and the participants could be grouped into a single group of participants, regardless of experience level. (*Table 7*)

3.4.2 Region Visual Parameter and Image Region Relationship

3.4.2.1 Time Before First Fixation in Region

There was a statistically significant difference between time before first fixation on a lateral cephalometric image and image region, regardless of experience level, *F* (4, 31) = 17.406, *p* < 0.001. For *R1*, there was an average of 2.411 seconds before first fixation. For *R2*, there was an average of 0.489 seconds before first fixation. For *R3*, there was an average of 2.660 seconds before first fixation. For *R4*, there was an average of 3.957 seconds before first fixation. For *R5*, there was an average of 10.335 seconds before first fixation. A statistically significant difference in time before first fixation was found between *R1* and *R2* (*p* (*R1, 2 (RegionEntry*)) = 0.015), and *R1* and *R5* (*p* (*R1, 5 (RegionEntry*)) < 0.001). There was no statistically significant difference in time before first fixation between *R1* and *R3* (*p* (*R1, 3 (RegionEntry*)) = 0.786) and *R1* and *R4* (*p* (*R1, 4 (RegionEntry*)) = 0.109). Thus the null hypotheses can be rejected. (*Table 7*, *Figure 18*)

3.4.2.2 Number of Fixations in Region

There was a statistically significant difference between number of fixations on a lateral cephalometric image and image region, regardless of experience level, *F* (4, 31) = 9.231, *p* < 0.001. For *R1*, there was an average of 42 fixations. For *R2*, there was an average of 27 fixations. For *R3*, there was an average of 33 fixations. For *R4*, there was an average of 25 fixations. For *R5*, there was an average of 17 fixations. A statistically significant difference in
number of fixations was found between R1 and R2 ($p_{R1,2} (\text{RegionFixation}) = 0.010$), R1 and R3 ($p_{R1,3} (\text{RegionFixation}) = 0.040$), R1 and R4 ($p_{R1,4} (\text{RegionFixation}) < 0.001$), and R1 and R5 ($p_{R1,5} (\text{RegionFixation}) < 0.001$). Thus the null hypotheses can be rejected. (Table 7, Figure 19)

### 3.4.2.3 Time Spent Viewing a Region

There was a statistically significant difference between time spent on a lateral cephalometric image and image region, regardless of experience level, $F (4, 31) = 8.883, p < 0.001$. For R1, there was an average of 16.682 seconds spent. For R2, there was an average of 8.547 seconds spent. For R3, there was an average of 11.809 seconds spent. For R4, there was an average of 9.231 seconds spent. For R5, there was an average of 5.801 seconds spent. A statistically significant difference in time spent was found between R1 and R2 ($p_{R1,2} (\text{RegionTimeSpent}) = 0.002$), R1 and R3 ($p_{R1,3} (\text{RegionTimeSpent}) = 0.023$), R1 and R4 ($p_{R1,4} (\text{RegionTimeSpent}) = 0.001$), and R1 and R5 ($p_{R1,5} (\text{RegionTimeSpent}) < 0.001$). Thus the null hypotheses can be rejected. (Table 7, Figure 20)

### 3.4.2.4 Number of Revisits to a Region

There was a statistically significant difference between number of revisits on a lateral cephalometric image and image region, regardless of experience level, $F (4, 31) = 9.409, p < 0.001$. For R1, there was an average of 15 revisits. For R2, there was an average of 9 revisits. For R3, there was an average of 12 revisits. For R4, there was an average of 16 revisits. For R5, there was an average of 6 revisits (SE = 1). A statistically significant difference in number of revisits exists between R1 and R2 ($p_{R1,2} (\text{RegionRevisits}) = 0.001$), R1 and R3 ($p_{R1,3} (\text{RegionRevisits}) = 0.014$), and R1 and R5 ($p_{R1,5} (\text{RegionRevisits}) < 0.001$). There was no statistically significant difference in number of revisits between R1 and R4 ($p_{R1,4} (\text{RegionRevisits}) = 0.375$). Thus the null hypotheses can be rejected. (Table 7, Figure 21)

### 3.4.3 Experience, Region Visual Parameter and Image Region Interaction

There was no statistically significant interaction between experience level, visual parameter, and image region in all instances examined ($p_{e, R1-5} (\text{RegionEntry}) = 0.396, p_{e, R1-5} (\text{RegionFixations}) = 0.537, p_{e, R1-5} (\text{RegionTimeSpent}) = 0.555$, and $p_{e, R1-5} (\text{RegionRevisits}) = 0.682$). Thus the null hypotheses cannot be rejected. (Table 7)
Table 7. Summary of quantitative results for parameter and image region.

<table>
<thead>
<tr>
<th>Viewing Parameter</th>
<th>Image Region</th>
<th>Participants</th>
<th></th>
<th></th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>e (vp)</td>
<td>i (vp)</td>
<td>e, i (vp)</td>
</tr>
<tr>
<td>Time before 1st Fixation (s)</td>
<td><strong>Alveolar Bone and Teeth</strong></td>
<td>2.411</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Region Sub-Classification</td>
<td>0.501</td>
<td>&lt;0.001</td>
<td>0.396</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Base of Skull and Calvarium</strong></td>
<td>0.489</td>
<td></td>
<td>0.015</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Upper and Middle Face</td>
<td>2.660</td>
<td></td>
<td>0.786</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lower Face</td>
<td>3.957</td>
<td></td>
<td>0.109</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cervical Spine, Airway and Area of the Neck</td>
<td>10.335</td>
<td></td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Number of Fixations</td>
<td><strong>Alveolar Bone and Teeth</strong></td>
<td>42</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Region Sub-Classification</td>
<td>0.360</td>
<td>&lt;0.001</td>
<td>0.537</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Base of Skull and Calvarium</td>
<td>27</td>
<td></td>
<td>0.010</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Upper and Middle Face</td>
<td>33</td>
<td></td>
<td>0.040</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lower Face</td>
<td>25</td>
<td></td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cervical Spine, Airway and Area of the Neck</td>
<td>17</td>
<td></td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Total Time (s)</td>
<td><strong>Alveolar Bone and Teeth</strong></td>
<td>16.682</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Region Sub-Classification</td>
<td>0.163</td>
<td>&lt;0.001</td>
<td>0.555</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Base of Skull and Calvarium</td>
<td>8.547</td>
<td></td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Upper and Middle Face</td>
<td>11.809</td>
<td></td>
<td>0.023</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lower Face</td>
<td>9.231</td>
<td></td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cervical Spine, Airway and Area of the Neck</td>
<td>5.801</td>
<td></td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Number of Revisits</td>
<td><strong>Alveolar Bone and Teeth</strong></td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Region Sub-Classification</td>
<td>0.418</td>
<td>&lt;0.001</td>
<td>0.682</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Base of Skull and Calvarium</td>
<td>9</td>
<td></td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Upper and Middle Face</td>
<td>12</td>
<td></td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lower Face</td>
<td>16</td>
<td></td>
<td>0.375</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cervical Spine, Airway and Area of the Neck</td>
<td>6</td>
<td></td>
<td>&lt;0.001</td>
<td></td>
</tr>
</tbody>
</table>

*Note:* All regions were compared to alveolar bone and teeth region. Values shown represent means. Statistically significant differences ($p<0.05$) are in red. $e = \text{experience}$, $vp = \text{visual parameter}$, $i = \text{image type}$. 
Figure 18. Mean time before first fixation (s) by image region as compared to R1.

![Time before First Fixation by Region](image)

**Note:** “∗” denotes a statistically significant difference ($p<0.05$). Bars connect image types that are statistically significantly different.

Figure 19. Mean number of fixations by image region as compared to R1.

![Number of Fixations by Region](image)

**Note:** “∗” denotes a statistically significant difference ($p<0.05$). Bars connect image types that are statistically significantly different.
**Figure 20.** Mean time spent examining (s) an image region as compared to R1.

![Diagram showing time examining by region](image)

*Note:* “*” denotes a statistically significant difference ($p<0.05$). Bars connect image types that are statistically significantly different.

**Figure 21.** Mean number of revisits by image region as compared to R1.

![Diagram showing number of revisits by region](image)

*Note:* “*” denotes a statistically significant difference ($p<0.05$). Bars connect image types that are statistically significantly different.
3.5 QUALITATIVE RESULTS (*Figures 22-29*)

3.5.1 Heat Map Analysis

The heat map showed that regions less commonly required for orthodontic treatment planning and diagnosis are not as readily examined and analyzed by orthodontists, regardless of experience. Instances in which orthodontists, regardless of experience, did not cover as much of an image as expected were instances where multiple abnormalities were present, and when variations of normal were suspected. (*Figures 22-25*)

**Figure 22.** Heat map of a normal lateral cephalometric image by experience level.

*Note:* *A* = original image; *B* = OR heat map; *C* = LEO heat map; *D* = MEO heat map. Image taken by Sarah Kaplan.
Figure 23. Heat map of an abnormal (obvious) lateral cephalometric image by experience level.

Note: $A =$ original image; $B =$ OR heat map; $C =$ LEO heat map; $D =$ MEO heat map. AOIs = enlarged sella turcica, pneumatized frontal sinus, frontal bossing, mandibular prognathism. Image taken by Sarah Kaplan.

Figure 24. Heat map of an abnormal (intermediate) lateral cephalometric image by experience level.

Note: $A =$ original image; $B =$ OR heat map; $C =$ LEO heat map; $D =$ MEO heat map. AOIs = severe high mandibular plane angle. Image taken by Sarah Kaplan.
Figure 25. Heat map of an abnormal (subtle) lateral cephalometric image by experience level.

Note: A = original image; B = OR heat map; C = LEO heat map; D = MEO heat map. AOIs = Hair-on-end appearance of skull, reduced diploic space, sella turcica bridging. Image taken by Sarah Kaplan.

3.5.2 Scan Path Analysis

Scan path analysis showed the increased distance covered on abnormal images, when compared to normal images. This distance was even further increased when obvious images were examined, as compared to normal image and abnormal images with intermediate or subtle abnormalities. Additionally, when comparing scan paths of a single participant over the course of multiple radiographs, it was clear that a similar and consistent search pattern approach was not followed, especially when an abnormality was detected, and that certain regions of a radiograph are excluded once an abnormality is detected. (Figure 26-29)
**Figure 26.** Scan path of a normal lateral cephalometric image by experience level.

*Note:* $A =$ original image; $B =$ OR scan path; $C =$ LEO scan path; $D =$ MEO scan path. Example of saccades (lines) and fixations (circles). Image taken by Sarah Kaplan.

**Figure 27.** Scan path of an abnormal (obvious) lateral cephalometric image by experience level.

*Note:* $A =$ original image; $B =$ OR scan path; $C =$ LEO scan path; $D =$ MEO scan path. AOIs = enlarged sella turcica, pneumatized frontal sinus, frontal bossing, mandibular prognathism. Example of saccades (lines) and fixations (circles). Image taken by Sarah Kaplan.
**Figure 28.** Scan path of an abnormal (intermediate) lateral cephalometric image, by experience level.

![Figure 28](image)

*Note:* A = original image; B = OR scan path; C = LEO scan path; D = MEO scan path. AOIs = severe high mandibular plane angle. Example of saccades (lines) and fixations (circles). Image taken by Sarah Kaplan.

**Figure 29.** Scan path of an abnormal (subtle) lateral cephalometric image by experience level.

![Figure 29](image)

*Note:* A = original image; B = OR scan path; C = LEO scan path; D = MEO scan path. AOIs = Hair-on-end appearance of skull, reduced diploic space, sella turcica bridging. Example of saccades (lines) and fixations (circles). Image taken by Sarah Kaplan.
CHAPTER 4

4 DISCUSSION

4.1 Experience Level and Overall, AOI and Region Viewing Parameter Quantitative Results Interpretation

An examination of the results showed that, for lateral cephalometric image interpretation, experience had no effect on the overall image viewing parameters (regardless of image type), viewing parameters of an AOI (regardless of abnormal sub-classification image type) and viewing parameters by image region (regardless of image region). This lack of effect suggests that orthodontists (OR, LEO and MEO) are similar in their capability of differentiating between normal and abnormal findings, AOIs, and regions, as well as levels of distraction by irrelevant visual data, and concentration and focus on the image, AOIs, and regions. These findings reveal that neither experienced orthodontists nor novice orthodontists demonstrate a greater visual efficiency when examining a lateral cephalometric image in whole, in part (looking at AOIs) and by region.

The sensorimotor expertise reversal effects through attention focusing may explain this unexpected result. When a novice is introduced to a new task, their lack of knowledge can be controlled through an overt step-by-step investigation that is slow and cognitively demanding. As expertise is developed through repeat practice, these step-by-step processes become implicit knowledge, which is fast and effortless. Unfortunately, execution of automatic processes, like evaluating a lateral cephalometric image, may be affected by requirement for step-by-step examination. For example, when experts work under conditions that require focused attention and execution in a step-wise manner, the expert automatic interpretation is disrupted, resulting in the regress to novice-like erratic evaluation and visual interpretation. On the other hand, novice performance is enhanced because their decreased knowledge of the task can be controlled by a stepwise execution through working memory.

Explicit monitoring theories can also explain these unexpected results. These theories explain that, for complex sensorimotor tasks that are completed under pressure and require explicit attention, expert’s real task performance is affected negatively. In other words, under pressure, experts become vulnerable to novice-like responses. This response is likely due to the increase in self-consciousness and anxiety of correct performance, which increases
attention paid to skill process and step-by-step control, and is called “choking under pressure.” In sum, results from these studies show that expert performance can be harmed by too much attention as a result of proceduralization and complexity of the task, and may explain the lack of differences in visual performance between ORs, LEOs, and MEOs in this study. Overall, these findings are not consistent with our hypotheses.

4.2 Overall Viewing Parameter and Image Type Interpretation

Experience aside, image type does have an effect on visual parameters, namely total distance covered and total number of blinks, explored in this study.

The distance traveled for abnormal images was increased when compared to normal images. This increase represents the heightened interest of orthodontists to perform a more thorough search to localize abnormalities when one is identified, in order to ensure complete interpretation of the radiograph. Furthermore, increased distance traveled in examination of an abnormal image indicates less efficient searching, typically as a result of an image layout. Lateral cephalometric radiographs may be considered complex radiographs. It is therefore expected that abnormal lateral cephalometric images would have a higher total distance covered than normal lateral cephalometric images. In sum, the type of search used by orthodontists is affected by the complexity and anatomical layout of the lateral cephalometric image. Interestingly, when image sub-classification was examined, no differences were noted; however, a trend in the data shows that there is increased distance covered on obvious images, followed by intermediate and, finally, subtle images. The more obvious the abnormality, the seemingly more complex the image, leading to an even larger total distance covered.

Additionally, there were more blinks for abnormal images than normal images, and more blinks for images with more obvious abnormalities than those with subtle abnormalities. This finding indicates a higher mental workload, requiring increased levels of concentration and focus by orthodontists when examining normal images and images with subtle abnormalities. This finding is supported by the results of a medical study that found that surgeons performing laparoscopic surgery self-reported fewer blinks and attributed fewer blinks to a higher level of mental workload needed for the surgery. As abnormalities become subtler, orthodontists must concentrate more deeply, and take less blinks; furthermore, normal radiographs must be distinguished as missing aberrations, thus orthodontists must concentrate on these images deeply, if not more deeply than images that do contain increasingly
subtle pathoses. In sum, the differences in visual parameters reflect cognitive processes that occur when reading lateral cephalometric images with abnormalities with varying levels of subtlety.

Although differences were found for the two aforementioned viewing parameters, there was no difference in total time searching, number of fixations, number and duration of saccades between normal and abnormal images. The lack of differences between image types indicates that there is no change in attention and focus, no change in degree of engagement and interest in areas of interest, no difference in amount of processing time required, no difference in amount of searching needed, and an inability to differentiate between areas of importance for image interpretation. Task complexity, time-on-task and task control may explain why these viewing parameters did not show differences between image types.

Participants were tasked with identifying and diagnosing abnormalities on lateral cephalometric images. There are four levels of task complexity (viewing, detection, decision, and problem-solving). Increasing task complexity occurs from viewing, to detection, to decision, to problem-solving. Problem-solving has many different paths to attain a desired outcome, may have several desired outcomes, and may require complex informational cues during task completion. Therefore, visual interpretation of each image type may have been affected by the complexity of the task, not the image itself.

Participants were also given control over each task, and were given unlimited time for image examination. As a result, ORs, LEOs, and MEOs had time to regulate visual and spatial processing demands from their working memories. However, because of the complexity of the lateral cephalometric image, both novices and experts experienced disorientation and extraneous information processing overload. Thus, the length of instruction may have affected the findings of this study, essentially resulting in a lack of significant differences in visual parameters, regardless of image type.

4.3. AOI Viewing Parameter and Abnormal Image Type Interpretation

Expertise aside, image type does have an effect on visual parameters when viewing an AOI.

There was faster time to first fixation in an AOI in images with obvious abnormalities than those with subtle abnormalities. This finding is expected, as images with obvious AOIs do have better attention-getting properties
than images with subtle AOIs. Furthermore, fast initial fixation on an abnormality may be indicative of a global search pattern by orthodontists.

There was also a greater number of fixations, number of revisits to, and time spent in AOIs of images with obvious abnormalities, and these viewing parameters decreased with increased AOI subtlety. First, increased number of fixations in obvious AOIs shows that orthodontists recognize these areas as important; Whereas in more subtle images, AOIs that may be significant, may not be recognized, or even seen as important, by orthodontists. Second, increased number of revisits to obvious AOIs is likely due to uncertainty, lack of understanding, and an attempt to rationalize and establish relationships and connections between AOIs, previous experience, and other radiographic findings by orthodontists of all experience levels. Lastly, increased amount of time searching AOIs on images with obvious abnormalities indicates higher attention and better focus than in AOIs on images with more subtle abnormalities.

In sum, lack of knowledge of normal and knowledge of features that signal abnormality may account for these findings. Obvious AOIs are examined longer, require more focus and attention, and have increased revisits than AOIs that are subtler. Subtle AOIs may go uninterpreted due to lack of visual analysis, likely because subtle AOIs lack visibility, and perhaps even meaningfulness, to orthodontists.

### 4.4 Region Viewing Parameter and Image Region Interpretation

Expertise aside, image region does have an effect on visual parameters when viewing lateral cephalometric images. Time to first fixation was fastest to the base of skull and calvarium (R2) and slowest to the cervical spine, airway and area of the neck (R5). This finding is unexpected, but is likely due to the anatomical layout and orthodontic analysis of a lateral cephalometric image. R2 does not have better attention getting properties than other regions of a lateral cephalometric image; however, R2 takes up the largest portion of a lateral cephalometric image, is centered on the computer screen, and contains an important point used in orthodontic analysis, sella. For these reasons, R2 may have been the region with quickest first fixation. R5 is located at the bottom of the computer screen and does not contain any orthodontic analysis points, planes or angles. For these reasons, R5 may have been the region with the most time before first fixation. The alveolar bone and teeth (R1), the upper middle face (R3) and the lower face
(R4) are similar in the amount of time to first fixation. This finding is also unexpected, and is also likely the result of
the anatomical layout and orthodontic analysis of a lateral cephalometric image. These three regions contain the
majority of orthodontic points, planes and angles used in orthodontic treatment planning and diagnosis and are
closely located to one another. In sum, this finding is surprising, as one would expect that dentists first examine the
alveolar bone and teeth. These results may indicate the use of a global-focal method of image interpretation by
orthodontists; however, it is most likely due to the layout of a lateral cephalometric image.

R1 had the greatest number of fixations, number of revisits (except when compared to R4), and time spent than other
lateral cephalometric regions. This finding is expected, as orthodontists, like dentists, do focus primarily on tooth-
bearing anatomy.  

R5 had the least number of fixations, number of revisits, and time spent than other lateral
cephalometric regions. This finding is thought-provoking as current orthodontic research has increased focus on
airway analysis, in particular for the role oropharyngeal and nasopharyngeal structures play in the development and
treatment of the dentofacial complex,  orthodontic expansion, sleep disordered breathing, surgery first
technique, and as a result of the development and increased use of cone-beam CT in orthodontic offices. These
are only some of the many ways airway analysis has intersected with orthodontics, and its ever-growing importance
in diagnosis and treatment planning.

Increased number of fixations in R1 shows that orthodontists recognize this region as important; whereas the other
regions may be significant, but may not be recognized, or even seen as important, by orthodontist. It is interesting
to note that R2 has increased fixations (although not more than R1 and R3) especially because lateral cephalometric
radiograph analyses do not focus on this region and because this region makes up the largest portion of a lateral
cephalometric image. It is also interesting that R4 has a low number of fixations, given the region’s increased
number of planes, angles and points used in orthodontic analysis and the region’s inclusion of the mandibular
symphysis, condyle, ramus, and body. These findings indicate that orthodontists have the greatest interest and
engagement in R1, followed by R2, R3, and R4, and least interest and engagement in R5.

An increased amount of time searching in R1 shows increased attention and focus on this region. This finding is
expected, and supports the fact that orthodontist pay most attention to tooth-bearing regions. Attention and focus
decrease from R1 to R3, followed by R4. Least attention and focus is spent on R2 and R5. This finding is expected,
as orthodontic analyses do not incorporate these areas in treatment planning and diagnosis, and so it is unlikely that orthodontist would pay attention to these regions.

Finally, an increased number of revisits to R1 indicate an attempt by orthodontists to rationalize and establish relationships and connections between regions. Orthodontists likely have an understanding of R1 given their educational background, but must relate their findings to previous experience and other radiographic findings in other regions. Furthermore, an increase in revisits to R1, R4, and R5, which contain the majority of orthodontic analyses points, planes and angles, indicate the use of these methods of analysis to form connections and relationships for image interpretation, and case diagnosis and treatment planning by orthodontists. Interestingly, R2 and R5 have the least number of revisits. This finding does not signify an orthodontist’s certainty and understanding of the regions. Rather, it signifies that orthodontists do not recognize these regions as important, and therefore they are less likely to return to R2 and R5 for repeat examination. Decreased revisits, here, indicates orthodontists’ inability to recognize an aberration from normal in R2 and R5 due to lack of knowledge and understanding of normal regional anatomy. Overall, R2 and R5 are more likely than other regions to go uninterpreted due to orthodontists lack of visual analysis, as a result of the regions’ seemingly lack of meaningfulness to orthodontic diagnosis and treatment planning. These regions are, in fact, important to orthodontists for diagnosis and treatment of each and every patient, and thus they must be interpreted just as R1, R3 and R4.

Overall, the results highlight an important aspect of visual search in lateral cephalometric interpretation: image layout. Instructional design is necessary for the visual flow and information processing to and within working memory. Two theories, Cognitive Theory of Multimedia Learning and Cognitive Load Theory, assume that there is a limitation to working memory, and that memory can be consumed by different loads, such as intrinsic loads (ie: task difficulty), extraneous loads (ie: layout of material) and germane load (ie: elaboration on the task). For learning or diagnosis to occur, information from visual inspection must be integrated with mental models (or schema) and prior knowledge. In this case, the more complicated an image (ie: lateral cephalometric image layout), the increased unnecessary extraneous load, resulting in increased intrinsic loads, thereby decreasing the working memory remaining that can foster efficient and effective cognitive processes for learning and understanding (or germane load). In order to mitigate the effects of increased extraneous and intrinsic loads, orthodontists tend to ignore everything they figured was not mandatory to solve the task (ie: How would you diagnose this case?). It is clear that
orthodontists do not just process lateral cephalometric images to view, detect, and decide; they must process the image in the context of the task (problem-solve), and therefore allocate their cognitive processes to do so efficiently and effectively. However, a positive relationship between image coverage and performance has been noted, and emphasizes the importance of complete radiograph interpretation.

4.5 Interaction of Experience, Viewing Parameter, and Image Type or AOI or Region

Quantitative Results Interpretation

This study found no interaction of experience, viewing parameter, and image type or AOI or region during lateral cephalometric image analysis by orthodontists indicating that orthodontists, regardless of experience, view and interpret lateral cephalometric images (in whole or in regions) similarly. Lack of interaction among the three main variables is likely the result of the complexity of lateral cephalometric images. Information-rich images can easily overwhelm learners, regardless of experience. Large amounts of irrelevant information challenge orthodontists to select relevant information, which is difficult. It is necessary, then, to approach expertise of information-rich images using methodological triangulation. Triangulation is defined as the “combination of two or more data sources, investigators, methodological approaches, theoretical perspectives, or analytical methods, within the same study.” While eye-tracking in this study tackled the perceptual aspects of visual expertise, follow-up questions for each image collected cognitive information to complete the data. This study, however, focused only on the analysis of the eye-tracking data.

Expertise reversal effects can also explain the performance of complex cognitive skills during lateral cephalometric interpretation by orthodontists. Expertise development mainly occurs through the growth of increasingly sophisticated cognitive schemas, or ways to store information in long-term memory. Expert performance is also not uniformly superior to novice performance. For instance, research in chess expertise show that experts did not perform superiorly to novices when trying to create meaningful combinations from random set-ups. This may indicate that impressing past knowledge on new images may not always be beneficial to performance. Several phenomena demonstrate deterioration in experts’ performance.

Deterioration in experts’ performance in interpreting lateral cephalometric images can be related to Cognitive Load Theory and intermediate effect in the Encapsulation Theory of medical expertise. Cognitive load theory focuses
on instructional design principles that are effective for novices, but detrimental for experienced learners, and (as discussed earlier) the limited capacity of working memory. Intermediate effect in the Encapsulation Theory of medical expertise focuses on recall of specific cases between novices and seasoned medical doctors. Advanced medical students could remember more and gave more elaborate explanations using biomedical knowledge and information, unlike experienced physicians. For experienced physicians, clinical knowledge and presentation played a larger role in diagnosis. However, seasoned physicians outperformed medical students on diagnostic performance, meaning they were more able to correctly identify pathology. Because no background medical or clinical information was provided to participants, it is possible that expertise reversal could have affected this outcome. Although seemingly different, these theories are similar in three ways. First, they are examples of ways experts and novices differ in learning and performance. Second, both theories have mechanisms that disrupt successful expert performance and enhance less experienced individual performance. Finally, they imply that different forms of activities or tasks may be appropriate for people with different levels of expertise.

Lack of knowledge and understanding of (and inability to recognize) abnormalities, increased susceptibility to distraction, decreased focus and attention, and inability to compare to contralateral anatomy, as a result of lateral cephalometric image layout and complexity, may result in decreased visual search efficiency and affect expertise. The results of this study beckons the creation of a systematic method of viewing lateral cephalometric images upon which orthodontists can form a strong visual search base.

### 4.6 Evaluation of Qualitative Results

#### 4.6.1 Heat Map Analysis

A heat map is a static map and does not give information about the order in which a subject views an image. The heat maps displayed little to no difference between orthodontists of various years of experience when looking at lateral cephalometric images. For example, it showed that ORs, LEOs, and MEOs have decreased fixations in some regions, and an overabundance in others. R2 was most likely to be overlooked, while R3 had the most fixations. Although it is not surprising that orthodontists tend to fixate on the dentition, which is very familiar to them, it is interesting to see that they heavily rely upon orthodontic landmarks in order to interpret an image. It is disappointing, however, that orthodontists do not focus as readily on those regions that are farther away from tooth-
bearing and orthodontic landmark regions, when these regions play an important role in the overall medical care orthodontists provide patients. This, again, is likely due to satisfaction of search and the attempt by orthodontists to reduce extraneous processing demands as a result of the task to diagnose (or problem solve) the lateral cephalometric image. \textsuperscript{61}

### 4.6.2 Scan Path Analysis

Scan paths describe where a subject looks and when they look there. Scan path analysis showed that orthodontists may use a global-focal search pattern when viewing lateral cephalometric images. \textsuperscript{62} A fast initial fixation on an AOI followed by a circumferential scan of the rest of the image, indicates use of this method. \textsuperscript{62} Furthermore, viewing the scan path of a single subject over multiple radiographs, clearly showed a different path for each image, especially if the image had an abnormality, and even more so when there were multiple abnormalities present in a single image. This finding is likely due to satisfaction of search, in which the orthodontist prematurely terminates his or her search, failing to look thoroughly at the unexamined portions of a radiograph once finding an abnormality. To some extent, orthodontists do continue to search a radiograph and cover the remaining areas of the image, before revisiting an abnormality for evaluation; however, excitement that comes with the detection of an abnormality, especially one that is outside the orthodontic analytical norms, can cause distraction and discontinuation of visual search. It is important to minimize the effects of satisfaction of search, as they can be detrimental to diagnosis and treatment planning. This finding highlights the need for strong emphasis on thorough search pattern when educating novices on how to read a lateral cephalometric image.

Finally, scan path analysis shows the emphasized use of landmark relationship by novices when they suspect an abnormality. Unlike panoramic images, where novices typically compare the contralateral side to the abnormal side, lateral cephalometric orthodontic analyses allow for the use of points, planes and angles to help in the evaluation of abnormalities. It is clear from this study that regions typically used for analytical reference are commonly revisited in order for orthodontists to make relationships and form an understanding of what aspects of the human anatomy present in the image deviate from normal development and growth. The heat maps generated, as well as the aforementioned complexities of lateral cephalometric radiographs and the ever important and educationally emphasized use of orthodontic landmarks, support this finding.
4.7 Study Limitations

Studies using eye tracking systems have several limitations. One such limitation for this study was the inability to match image-viewing conditions with the orthodontic office setting. For instance, the inability to adjust lighting in the room may have affected the ability of a participant to detect small and low contrast abnormalities. Additionally, it was impossible to adjust contrast and brightness of the image using the program software. To adjust for these issues, all images were acquired digitally, to avoid information loss that may occur with film-based imaging. Furthermore, limitations in range of head movement beyond 50-75 cm made it impossible to collect data that matches the dynamic orthodontic office environment. To minimize this effect, a stabilization-free eye tracker was selected.

Additionally, an eye tracker cannot appreciate a subject’s peripheral view. It is possible for participants to have visualized an AOI, despite having fixed outside of the AOI, using their peripheral vision. One study found that subject who may have never directly viewed the periphery of the cranium, were still able to identify it, and that this was especially true for areas with increased contrast differences. 43 Eye trackers track only what is present in an individual’s foveal vision, which accounts for less than 8% of the visual field. 35 The remaining 92% of the visual field is composed of parafoveal and peripheral vision. 35 These regions of vision give information about what is happening around us, and a general idea of color, shape and motion. 35 Although details cannot be detected with parafoveal and peripheral vision, they add an important layer to our global impression of what is taking place outside of our foveal vision. 35 This makes eye-tracking data tricky, since although a participant may not fixation on a region, they may be aware of what is there. 35

Furthermore, interpreting mapped fixations, and scan paths, developed by the eye tracking system can be deceiving. Often, a fixation is registered, although a user may not have cognitively registered the view in their brain. 35 Over-reporting fixations inaccurately reports attention. Eye trackers are not mind-reading devices – They can only tell us what a person looks at, and not why. Little is still known about the cognitive processes that can be deduced from eye movement. 35 The observed eye movements reflect many ongoing cognitive processes. 35 Thus, it is prudent to avoid making inferences from these movements about what process is going on. 35 Often, the explanation of eye movements is an idea, not a conclusion. 35 This limitation can be improved by adding follow-up questions that are
rooted in theory, providing subjects an appropriate experimental task, and asking for verbal, or think-aloud, data for explanation of why someone looked somewhere, previously introduced as triangulation.\textsuperscript{34,35,68,69}

Moreover, although this study did collect information using triangulation methodology, it could have been improved. It is possible that the differences in the intersection of experience, viewing parameter, and image type or AOI or region lie in the cognitive information, which was not explored in this study. This study also did not link the analysis of reported cognitive information with eye tracking data. Furthermore, verbal data would have been good to improve interpretation of scan pathways with the understanding of “why” one looked where they did.\textsuperscript{56,68,69}

In addition, eye-tracking systems may have difficulty calibrating participants with droopy eyelids (perhaps from anatomy, sleeping or viewing fatigue), contact lenses, narrow glasses, wearing mascara or of Asian descent. An earlier study found that Asian participants had decreased trackability, accuracy and precision when compared to African and Caucasian subjects.\textsuperscript{51} However, in this study, participant ethnicity was not taken into account if appropriate calibration could be obtained in order to improve sample size.

Furthermore, a limitation of eye tracking systems is the inability to generalize findings. Eye tracking studies, like this one, typically have a decreased sample size due to the time-consuming process required for data collection.\textsuperscript{61} Although this study was similar to previously published studies in terms of number of subjects in each group, a small number of participants in each group could have resulted in sampling error.\textsuperscript{61}

We did not find significant results at 20 seconds between each group. While power may be a limitation of the study, a post-hoc power analysis indicated that if there were two groups (OR and EO), a total sample of 400 participants (200 participants per group) would be needed for the difference we found to be statistically different. The two standard curves overlap by a third of a standard deviation. From a practical standpoint, a third of a standard deviation may not be practically significant. Bigger differences were expected based upon literature review and earlier published studied; however, this study did not have those same findings.

Finally, an inherent limitation of studies using an eye tracker is the participant’s knowledge of an eye tracker presence, also known as the Hawthorne Effect. The Hawthorne Effect is as “the stimulation to output or accomplishment that results from the mere fact of being under observation.”\textsuperscript{36} In general, participant expected each radiograph to contain an abnormality, despite having been told that there were a certain number of images free of
abnormality (i.e. “normal”) that had been pre-selected by a focus group. As a result, participants were likely to conduct a more thorough search than they typically would if not under scrutiny. This finding likely affected the time taken to evaluate each image and the number of fixations in different regions investigated. Additionally, experienced orthodontists may have “choked under pressure” and thus the results may be unintentionally affected. Although this may be a limitation of the study, it may be a novel finding, that every time an orthodontist (regardless of experience level) reviews a new lateral cephalometric image, they revert to the step-wise examination process used by novices as a result of increased focus and attention.

Beyond those related to the use of an eye tracker, the study had other limitations. First, this study excluded patient medical and dental history. Medical and dental history is critical in making each image clinically relevant. However, this information was eliminated as it could introduce bias in the search pattern and approach used by participants. An earlier study demonstrated that, during the interpretation of electrocardiograms, participant diagnosis was influenced by information regarding subject age, sex and profession. Unfortunately, it has been reported that this additional information could help to distinguish experienced clinicians from novices. Second, all resident class-years were grouped together into an “orthodontic residents” group. It would be interesting to examine if there are differences between class years; however, the sample size in this study was too small for each residency group to perform this investigation.

Finally, a limitation of this study was the assumption that the data collected was normally distributed. Fortunately, a two-way repeated measures ANOVA is a robust mathematical evaluation of data. Normality was assumed because there is no non-parametric alternative for a two-way repeated measures ANOVA. As a result, there is a possible increase in Type 1 Error, or chance of higher number of false positive reporting in the study. Lowering statistical significance of this study to \( p > 0.01 \), would result in the drop out of 7 significant finding overall, and lowering statistical significance of this study to \( p > 0.005 \), would result in the drop out of 12 significant finding overall. This indicates that it is likely like the data is not normal and some false positives have been reported in this study.
4.8 Implications for Orthodontic Education and Practice

Abnormality localization and interpretation are two equally important components of orthodontic treatment planning and diagnosis. Localization, as mentioned previously, relies upon knowledge of human anatomy, its variants, and a strong search algorithm. This study implies the importance of a systematic instructional design and training for novices from which an organic viewing method can develop. Gaze modeling can help increase motivation, interest, and engagement in radiograph interpretation, as well as model for what to focus on, in what order to focus, and reasoning for novices. 61

Instructional methods that are beneficial for novices are not beneficial for experts or more knowledgeable students. 56 Although orthodontic residents enter into their post-graduate education with a background in radiology, human anatomy, pathology, and other proceduralized tasks and automatic skills for evaluating other types of radiographic images necessary for general dentistry, such as panoramic images, periapical and bite wing images, instructional guidelines for the learning of new tasks as they relate to lateral cephalometric images should be developed. However, because redundancy of information is detrimental to experienced practitioners, it is important to tailor information format and guidance required. 56 As a result, teaching procedures, methods, and systems early in an orthodontic curricula would be beneficial to the novice, who could create overlaps between their learned and proceduralized knowledge and provided educational guidance. 56 Furthermore, if differences in radiographic interpretation between novice and experienced orthodontists really do not exist as shown in this paper, it is even more important to emphasize the significance of a standardized search pattern approach, in order to assure full, complete, and thorough search of all structures visible on a lateral cephalometric image, as well as the wide variety of normal anatomy.

Furthermore, a deep understanding of the etiology of pathoses is key to image interpretation, not experience alone. This means, that in addition to teaching a method of image search, it is important to also teach the pathophysiology of disease and abnormality development, so as to underpin the reasons why a particular feature is seen. 11 One study found that a deeper understanding led to improved performance by experts in observing radiologist search of lung nodules on a PA chest radiograph. 31 This finding reinforces the need for a deep understanding of pathophysiology,
in addition to increased exposure of students to lateral cephalometric images, and the development of a search pattern for lateral cephalometric images.

Together, these findings can be used in orthodontic education to improve the way lateral cephalometric image interpretation is taught. The main features of this study show that novices and experienced orthodontists view lateral cephalometric images similarly over the course of their professional lives, that subtlety of an abnormal may have an effect in how the image is viewed, and that although a full image is covered, majority of attention is paid to regions of a lateral cephalometric image that are important to orthodontic diagnosis and treatment planning. At present, orthodontic residencies tend to teach orthodontic analyses, which seem to dictate image interpretation; however, these analyses do not cover the full radiographic image, thus leaving regions of the image unexamined by orthodontists. Thus, the use of a systematic search method of lateral cephalometric images or the creation of a short video to be reviewed by residents to remind them to look systematically at all lateral cephalometric images is something that may be beneficial to the field of orthodontics, so as to create a subconscious habit for orthodontists to always perform the same search strategy each time for the course of their professional career to decrease the chance of visual error, potentially increase the number of abnormalities localized, and eventually lead to increased efficiency in localization and diagnosis of pathology.

Eye-movement modeling examples, or EMMEs, can effectively guide visual attention in order to teach clinical reasoning skills. Videos of how to effectively and appropriately view images in order to complete tasks (like problem solving), in conjunction with discussion and verbal records of relevant areas and methods of interpretation, can help direct novices. To be effective, EMMEs must encourage learning from example, allow the cognitive and perceptual processes to be made available to novices, and provide guidance for visual input processes. For instance, videos can use a spotlight of eye movement superimposed on an image guiding a viewer’s attention and focus with an audio over recording of expert verbal explanation of interpretation. It is possible that a video of this kind could benefit orthodontic education, ensure full evaluation of a lateral cephalometric image, and improve diagnosis and treatment planning in the field.
4.9 Future Directions

There are many possible direction future directions that can be adapted from this study; however, no matter what direction is taken, design, selection and task sequencing must be carefully tailored to knowledge and experience level.

First, the use of a cone-beam computed tomography (CBCT) images are becoming mainstream in the field of orthodontic. A future study could evaluate the search pattern used by experts when evaluating these types of images. Information collected from such a study could be used to create a search pattern for evaluation of CBCTs, ensuring a higher standard of care for patients and more accurate interpretation of CBCT volumes.

Second, since this was the first observational study of its kind in orthodontics, a multitude of questions were asked of participants that, in the end, were not used in the analysis. These questions were used for study triangulation and may provide insight into the diagnostic accuracy of orthodontists. It is possible that differences between ORs, LEOs, and MEOs exist only in the diagnostic accuracy portion of lateral cephalometric images, and this information could also help determine if participants primarily have search error, recognition error, or decision error. Second, it would also be interesting to see if there are differences between male and female participants and between residents at different programs. This would require a much larger number of participants with similar level of experience and program location. Third, data evaluating participant confidence in image interpretation was collected. There may be a link between confidence level and search strategy that could be answered in future studies. Finally, it would be interesting to repeat a similar study using oral maxillofacial radiologists in place or in addition to MEOs.

Third, the study could be repeated with some changes to the design. For instance, images could be repeated within the sample to see if there are differences in time, accuracy, scan path, etc. between images early and late in the sample. This may help mitigate any affect of viewing fatigue.

Fourth, this study could be performed with a larger sample to investigate for any potential differences between resident class years. It is possible there are differences within the OR group. Understanding differences between residency class years may be an interesting in tracking interpretation changes throughout post-graduate orthodontics education.
Finally, this study was setup in such a way that the subjects were not provided assistance during the procedure. An eye-tracker as a fantastic instrument that gives access into the minds of students and provide them with instant feedback. Although employing this in all schools would be impractical for a predoctoral program, it could certainly be feasible for a residency graduate program, like on in orthodontics, where the ratio of professors to students is much lower.
CHAPTER 5

5 CONCLUSION

This study found, no significant difference in:

1. The effect of experience level (OR, LEO, MEO) on viewing parameter of an image overall (total time spent viewing, total number of fixations, total distance covered, total number of blinks, total number and length of saccades), an AOI (time before first fixation, number of fixations, time spent viewing, number of revisits), and region [R1 (alveolar bone and teeth), R2 (calvarium and base of skull), R3 (upper and middle face), R4 (lower face), and R5 (cervical spine, airway and area of the neck)].

2. The interaction of experience, viewing parameter (as listed above for image overall, AOI or region), and image type (n, a, o, i, s).

There was a significant difference in:

1. The effect of image type (n, a, o, i, s) on viewing parameter for:
   a. Total distance covered (n, a)
   b. Total number of blinks (n, a, o)

2. The effect of image type (o, i, s) on AOI viewing parameter for:
   a. Time before first fixation (o, s)
   b. Number of fixations (o, i, s)
   c. Time spent viewing (o, i, s)
   d. Number of revisits (o, i, s)

3. The effect of image type (o, i, s) on region viewing parameter for:
   a. Time before first fixation (R1, R2, R5)
   b. Number of fixations (R1-R5)
   c. Time spent viewing (R1-R5)
   d. Number of revisits (R1, R2, R3, R5)

These findings indicate that experience does not affect overall examination performance of lateral cephalometric images, AOIs, and region, and there are no interactions affecting examination performance between the experience
and image, AOI and region. Experience aside, different levels of image and AOI subtlety and lateral cephalometric
image region do affect examination performance.

The ultimate goal of eye-tracking research in medicine and dentistry is to improve image interpretation by avoiding
errors in visual search. Identification of strategies that may improve clinician performance and treatment can come
from an understanding of the perceptual process of image visualization. These strategies are important to integrate
into orthodontic radiology training.
APPENDIX

Appendix A. SUNY University at Buffalo Institutional Review Board Approval Letter

University at Buffalo Institutional Review Board (UBIRB)
Office of Research Compliance | Clinical and Translational Research Center Room 5018
875 Ellicott St. | Buffalo, NY 14203
UB Federallywide Assurance ID#: FWA00008824

APPROVAL OF SUBMISSION

March 9, 2017

Dear Sarah Kaplan:

On 3/9/2017, the IRB reviewed the following submission:

<table>
<thead>
<tr>
<th>Type of Review:</th>
<th>Initial Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title of Study:</td>
<td>From Novice to Expert: A Comparison of Examination Performance of Lateral Cephalometric Images by Orthodontic Residents and Experienced Orthodontists</td>
</tr>
<tr>
<td>Investigator:</td>
<td>Sarah Kaplan</td>
</tr>
<tr>
<td>IRB ID:</td>
<td>STUDY00001212</td>
</tr>
<tr>
<td>Documents Reviewed:</td>
<td>• Supplemental Information-2 Specs: RED250mobile, Category: Other; • HIPAA Waiver (HRP-612), Category: Other; • Main Study Adult Consent to Participate in a Research Study, Category: Consent Form; • Sarah Kaplan IRB Protocol, Category: IRB Protocol; • Supplemental Information-2 Specs: Red-n, Category: Other; • Supplemental Information-2 Specs: RED-m, Category: Other; • Focus Group Adult Consent to Participate in a Research Study, Category: Consent Form; • Supplemental Information 1, Category: Recruitment Materials</td>
</tr>
</tbody>
</table>

The Initial Study materials for the project referenced above were reviewed and approved by the SUNY University at Buffalo IRB (UBIRB) by Expedited Review.
Appendix B. Google Form survey used for first round of radiograph selection

Final Radiograph Selection

As a reminder: The subcategories (subtle, intermediate, and obvious) refer to the level of difficulty to detect the abnormality, NOT to diagnose it.

Page 13 *

- Reject
- Normal
- Abnormal - Subtle
- Abnormal - Intermediate
- Abnormal - Obvious

Please name the abnormality present and its associated localization (A = base of skull, B = upper and middle face, C = lower face, D = cervical spine, airway and neck, E = alveolar bone and teeth). There may be more than one abnormality present in the radiograph. For example, "sella bridging, A; impacted canine, E." If the image is "rejected" please type the reason for rejection.

Short answer text

........................................................................................................................................

........................................................................................................................................
Appendix C. Participant informed consent.

Permission to Take Part in a Human Research Study

University at Buffalo Institutional Review Board (UBIRB)

University at Buffalo Institutional Review Board (UBIRB)
Office of Research Compliance | Clinical and Translational Research Center Room 5018
875 Ellicott St. | Buffalo, NY 14203
UB Federalwide Assurance ID#: FWA00008824

Adult Consent to Participate in a Research Study

Title of research study: From Novice to Expert: A Comparison of Examination Performance of Lateral Cephalometric Images by Orthodontic Residents and Experienced Orthodontists

Version Date: 2

Investigator: Sarah Kaplan

Why am I being invited to take part in a research study?
You are being invited to take part in a research study because you are a current orthodontic resident or experienced orthodontists with 5-10- or 25-years post-residency.

What should I know about a research study?
- Someone will explain this research study to you.
- Whether or not you take part is up to you.
- You can choose not to take part.
- You can agree to take part and later change your mind.
- Your decision will not be held against you.
- You can ask all the questions you want before you decide.

Who can I talk to?
If you have questions, concerns, or complaints, or think the research has hurt you, talk to the research team at sarahkap@buffalo.edu. You may also contact the research participant advocate at 716-888-4845 or researchadvocate@buffalo.edu.

This research has been reviewed and approved by an Institutional Review Board (“IRB”). You may talk to them at (716) 888-4888 or email ub-irb@buffalo.edu if:
- You have questions about your rights as a participant in this research
- Your questions, concerns, or complaints are not being answered by the research team.
- You cannot reach the research team.
- You want to talk to someone besides the research team.
- You want to get information or provide input about this research.
Permission to Take Part in a Human Research Study

Why is this research being done?
The purpose of this study is to evaluate and compare the search patterns used by orthodontic residents and experienced orthodontists when interpreting lateral cephalometric radiographs.

How long will the research last?
We expect that you will be in this research study for 45 minutes.

How many people will be studied?
We expect about 120 people in this research study.

What happens if I say yes, I want to be in this research?
If you want to be in this research, the primary investigator will set up a time with you to perform the experiment. The experiment itself is expected to take between 30 and 45 minutes, and will vary depending on how much time you spend searching the radiographs. You will be responsible to answer questions to collected demographic data, then view twenty (20) lateral cephalometric radiographs and answer pertaining questions to each radiograph. An eye tracking device will be used to follow your eye movements. You should search these radiographs for any abnormality (excluding dental caries and periodontal disease), like you would do in a normal clinical setting. Do not change your search pattern for this experiment. After each radiograph, you will be responsible for completing a chart relating to your findings on this radiograph. The principal investigator or the supervisor will always be present during these experiments if any problems or questions arise.

What are my responsibilities if I take part in this research?
If you take part in this research, you will be responsible to complete the full visual search presentation at the orthodontic meeting of your selection.

What happens if I do not want to be in this research?
Your participation in this research study is voluntary. You may choose not to enroll in this study.

What happens if I say yes, but I change my mind later?
Participation in this research project is voluntary. You can leave the research at any time it will not be held against you. Data that is collected to the point of withdrawal will be used for analysis. You will be asked to explain the extent of their withdrawal.

Is there any way being in this study could be bad for me?
There are minimal risks associated with these procedures.

- Physical risks: soreness from prolonged sitting (minimal risk), visual fatigue (minimal risk) and eye injury (minimal risk).

Will being in this study help me in any way?
Permission to Take Part in a Human Research Study

We cannot promise any benefits to you or others from your taking part in this research. However, possible benefits include requesting the results of your visual scan test to learn more about your own visual search pattern. Results of the study may lead to improvement in your professional field of orthodontics, specifically in diagnosis and treatment planning methodology.

What happens to the information collected for the research?
Efforts will be made to limit the use and disclosure of your personal information, including research study and medical or education records, to people who have a need to review this information. We cannot promise complete secrecy. Organizations that may inspect and copy your information include the IRB and other representatives of this organization. The identities of the participants will be kept confidential. Participant data will be anonymized as will information related to the images being displayed. The identities and email address of the participants will be known only to the principal investigator.

What else do I need to know?
You will not be paid for participating in this study.

Signature Block for Capable Adult
Your signature documents your permission to take part in this research. By signing this form you are not waiving any of your legal rights, including the right to seek compensation for injury related to negligence or misconduct of those involved in the research.

___________________________  ______________________
Signature of subject              Date

___________________________
Printed name of subject

___________________________  ______________________
Signature of person obtaining consent       Date

___________________________
Printed name of person obtaining consent
REFERENCES


5. Interpretation Merriam-Webster: Merriam-Webster.


35. Bergstrom JR, Schall AJ. Eye tracking in user experience design. Amsterdam: Elsevier/ Morgan Kaufmann; 2014.


