The Influence of Experience and Training in the Examination Pattern of Panoramic Radiographs: Results of an Eye Tracking Study

by

Daniel P. Turgeon

A thesis submitted in conformity with the requirements for the degree of Master of Science in Oral and Maxillofacial Radiology
Discipline of Oral and Maxillofacial Radiology
University of Toronto

© Copyright by Daniel P. Turgeon 2014
The Influence of Experience and Training in the Examination Pattern of Panoramic Radiographs: Results of an Eye Tracking Study

Daniel P. Turgeon

Master of Science in Oral and Maxillofacial Radiology

Discipline of Oral and Maxillofacial Radiology, Faculty of Dentistry
University of Toronto

2014

Abstract

Physician training has greatly benefitted from insights gained in understanding the manner in which experienced practitioners search medical images for abnormalities. The objective of this study is to compare the search patterns of dental students and certified oral and maxillofacial radiologists (OMRs) over panoramic images. An eye tracking system was used to accrue multiple parameters for both groups searching for abnormalities on 20 panoramic radiographs. Compared with students, OMRs displayed more consistent search patterns, and spent overall less time with fewer blinks, saccades and eye fixations. As students frequently changed their search patterns between different images, undergraduate dental education programs should emphasize the need for developing systematic search strategies. Moreover, as students were often distracted by image artifacts, greater emphasis should be placed on an understanding of panoramic image artifact generation so that they are less distracted by such areas, and develop greater focus on areas of relevance.
Acknowledgments

Thank you to my supervisor and mentor, Dr. Ernest Lam, for allowing me to work with such an interesting research subject. At the same time providing me with an outstanding education in Oral and Maxillofacial Radiology, all the while improving my English skills. Your teaching, even though intense sometimes, showed me to be meticulous and rigorous in my work.

Thank you Dr. Michael Pharoah for showing me how to excel in the interpretation art, but also how to accept being wrong and using that to become a better diagnostician. I will do my best to apply this, wherever life takes me.

Special thanks to Drs. Susanne Perschbacher, Michael Pharoah, and Ernest Lam for the time you spent analyzing hundreds of radiographs.

Another special thanks to my co-resident, present and past, for acting as guinea pigs when I needed your help.

Un merci special à ma co-résidente et future collègue, Catherine. Certaines périodes de ces trois dernières années ont été plus difficiles que d’autres, mais nous y sommes tout de même parvenus!

Finalement, un dernier mot à celle qui m’a encouragé pendant ces trois (longues) années. Malgré la distance qui nous séparait, ton appui s’est fait sentir jusqu’ici et m’a permis de supporter cet éloignement. De tout mon coeur, merci To Nhu!
# Table of Contents

## Contents

Acknowledgments.................................................................................................................... iii

Table of Contents.................................................................................................................. iv

List of Tables .......................................................................................................................... vii

List of Figures .......................................................................................................................... viii

List of Appendices ................................................................................................................... x

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

1.1 General Considerations .................................................................................................. 1

1.2 Localization of Pathoses .............................................................................................. 1

1.3 Strategies in Image Interpretation ................................................................................. 2

1.4 Eye Tracking ................................................................................................................... 4

1.4.1 General Considerations ............................................................................................. 4

1.4.2 Dentistry ...................................................................................................................... 8

1.4.3 Data Output .................................................................................................................. 8

1.5 Panoramic Radiography ............................................................................................... 9

1.6 Summary ......................................................................................................................... 10

1.7 Aim and Statement of Problem .................................................................................... 10

1.8 Objectives ....................................................................................................................... 11

1.9 Hypotheses ..................................................................................................................... 11

1.9.1 Alternate Hypotheses ............................................................................................... 11

1.9.2 Null Hypotheses ........................................................................................................ 12

2 Materials and Methods ..................................................................................................... 14

2.1 Observer Selection ......................................................................................................... 14

2.1.1 Sample Size ............................................................................................................... 15
2.2 Selection of the Panoramic Radiographs .......................................................... 16

2.3 Eye Tracking System ...................................................................................... 18

2.3.1 Software Programming .............................................................................. 19

2.4 Experimental Protocol ................................................................................... 19

2.5 Statistical Analysis ......................................................................................... 20

3 Results ............................................................................................................... 24

3.1 Quantitative Results ....................................................................................... 24

3.1.1 Total Time Examining the Radiograph ....................................................... 24

3.1.2 Number of Fixations on the Radiograph .................................................... 25

3.1.3 Distance Covered on the Radiograph ......................................................... 25

3.1.4 Number of Blinks on the Radiograph ....................................................... 26

3.1.5 Number of Saccades on the Radiograph ................................................... 26

3.1.6 Length of Saccades on the Radiograph ..................................................... 26

3.1.7 Time Before First Fixation in Area of Interest (AOI) ................................. 27

3.1.8 Number of Fixations in AOI ................................................................. 27

3.1.9 Total Time Looking an AOI ................................................................. 27

3.1.10 Number of Revisits in AOI ................................................................. 28

3.2 Descriptive Results ....................................................................................... 28

3.2.1 Heat Maps ............................................................................................... 28

3.2.2 Scan Paths .............................................................................................. 29

4 Discussion ......................................................................................................... 35

4.1 Evaluation of the Quantitative Results ......................................................... 35

4.1.1 Parameters Involving the Radiograph ...................................................... 35

4.1.2 Parameters Involving the Area of Interest (AOI) ....................................... 36

4.2 Evaluation of the Descriptive Results .......................................................... 36

4.3 Study Limitations ......................................................................................... 38
4.4 Implications for Oral Radiology ........................................................................... 39
4.5 Future Directions ................................................................................................. 41
5 Conclusion .............................................................................................................. 44
References .................................................................................................................. 45
List of Tables

Table 1. Summary of the quantitative results (mean ± SD). .................................................. 30
List of Figures

**Figure 1.** Scan path of a dental student of a normal panoramic radiograph. Example of saccades (lines) and fixations (circles). ................................................................. 13

**Figure 2.** Screenshot of the Experiment Center showing the group randomization. .............. 21

**Figure 3.** Points of the calibration screen.................................................................................. 22

**Figure 4.** Photograph of the setup. The participant is to the left, the observer is to the right. ..... 22

**Figure 5.** Photograph of the setup (close-up). ............................................................................ 23

**Figure 6.** Results of calibration. .................................................................................................. 23

**Figure 7.** Heat map of the dental students of an abnormal (intermediate) panoramic radiograph. There is a supernumerary tooth distal to the maxillary left third molar (2.8)................................. 31

**Figure 8.** Heat map of the OMRs of an abnormal (intermediate) panoramic radiograph. There is a supernumerary tooth distal to the maxillary left third molar (2.8)................................. 31

**Figure 9.** Heat map of the dental students of a normal panoramic radiograph. ....................... 32

**Figure 10.** Heat map of the OMRs of a normal panoramic radiograph. ................................. 32

**Figure 11.** Heat map of the dental students of an abnormal (intermediate) panoramic radiograph. There is a benign tumour (hemangioma) located distal to the maxillary right second molar (1.7). .............................................................................................................. 33

**Figure 12.** Heat map of the OMRs of an abnormal (intermediate) panoramic radiograph. There is a benign tumour (hemangioma) located distal to the maxillary right second molar (1.7). ........ 33

**Figure 13.** Heat map of the dental students of an abnormal (obvious) panoramic radiograph with multiple abnormalities. There are four keratocystic odontogenic tumours (KOTs) located in the following regions: maxillary left molar, mandibular left molar, mandibular left first premolar, mandibular right canine. .............................................................................................................. 34
Figure 14. Diagram of a panoramic radiograph for localization of abnormality................. 54
List of Appendices

Appendix A. Health Sciences Research Ethic Board Approval Letter............................................. 49
Appendix B. Email sent to dental students for participation in the study........................................... 50
Appendix C. Email sent to OMRs through “oradlist” for participation in the study......................... 52
Appendix D. Instruction and example for the panoramic radiographs selection........................... 53
Appendix E. Participant informed consent ....................................................................................... 55
1 Introduction

Radiography plays an important role in the diagnosis and treatment planning of any patient, whether in the medical or dental domain. The most commonly used radiographs in dentistry are periapical, bitewing and panoramic radiographs, and since these are two-dimensional representations of the three-dimensional anatomy, image interpretation can be difficult. It is clear that some practitioners are better able to interpret radiographs than others, but how do they do it? Multiple studies in the medical radiology field have been performed over the years to answer that question by analyzing the patterns of search used by radiologists and medical radiology residents. These studies have led to a better understanding of the ways in which radiology specialists see abnormalities on radiographs. These findings then have the potential to be used to create different methods of teaching radiologic feature identification and interpretation, using the experienced practitioners' method of analysis. In dentistry, little work has been done in this area.

1.1 General Considerations

Interpretation is defined as the "act or results of explaining something" [1]. Even though this definition is not specific for radiologic interpretation, it still shows the process radiologists use on a daily basis; putting into clear and articulate sentences their observations and developing an understanding of what these observations mean. Although "picture matching" is sometimes used as a means of radiographic interpretation, it soon becomes obvious for every practitioner that the range of appearances for every type of abnormality is too broad to rely solely on this technique, especially early in a career. Fortunately, other strategies have been developed over the years.

1.2 Localization of Pathoses

In contrast to medicine where radiologists are responsible for making and interpreting the images prescribed by the medical community at large, dentists and dental specialists are usually responsible for making and interpreting their own images. One method that is currently taught to dental students for panoramic radiographs in particular, is the "region-by-region" systematic
approach. A version of this method is described by White and Pharoah \(^7\). This strategy divides a panoramic radiograph into multiple regions: mandible, midface, soft tissues, superimpositions and ghost images and finally, the dentition. This method directs the untrained eyes of novice clinicians (i.e. dental students) to review specific regions and features of the panoramic radiograph. Recently, Khalifa \(^{16}\) showed that training dental students to use a systematic search strategy led to an improvement in the identification of pathoses on panoramic radiographs. Unfortunately, her work also found an increased frequency of false-positive observations. This infers that training to localize an abnormality and training to identify it correctly are two separate tasks, but both are necessary to make an adequate radiologic diagnosis. It is unclear if experienced practitioners search panoramic radiographs in this systematic way although recent studies suggest they actually use a free search pattern \(^7,^{33}\), but with a high level of accuracy.

While the study by Khalifa \(^{16}\) provided some interesting results, there were some limitations. First, there was no proof that the participants actually followed the instructions and applied the systematic search strategy to the radiographs. There is therefore a possibility that the differences seen between the groups could simply be due to the review of panoramic radiography provided by the instructional video, rather than the use of the systematic search strategy that was taught. As the author points out, the use of a device that could track the eye movements of participants could confirm the search strategy of participants. Another major point the author raises is that the lack of prior knowledge in normal anatomy and pathology, especially in the students group, could have led to their over-reporting of false-positives. This could have been corrected for by adding a component about normal anatomy and artifacts in the video presentation. Lastly, the absence of patient information and clinical history could have been detrimental to the diagnostician.

1.3 Strategies in Image Interpretation

Baghdady et al. \(^7,^{23}\) recently proposed two approaches to image interpretation used in oral radiology. The first approach is the analytical strategy; sometimes referred to as “forward reasoning”. This strategy relies on a step-by-step analysis of the features of a radiographic finding, and eventually, a diagnosis is made, based on these features. The proposed analytical method described by White and Pharoah \(^7\) directs clinicians to identify the following features of an abnormality: location, size and shape, borders, internal structure, and effects on the
surrounding structures. Once a feature has been identified, the clinician needs to decide if it represents something normal or abnormal. If a feature is deemed to be abnormal, then is the abnormality acquired or developmental? And finally, if an abnormality is identified, in which category of pathosis (cyst, benign or malignant neoplasm, inflammation, bone dysplasia, vascular, metabolic, or trauma) does it fit best? Obviously, the link that is made between the identified feature and the category of pathosis must come from a strong understanding of a pathology’s radiographic features. This analytical strategy has been used to teach novice clinicians how to interpret radiographs, and to avoid bias and premature conclusion-making in the analytical process.

The second approach is the non-analytical strategy, sometimes referred to as “backward reasoning”. This is the strategy that is predominantly used by experienced radiologists. This strategy assumes that viewing the abnormality as a whole rather than its individual radiographic features leads to a provisional diagnosis. This approach is often considered to be done automatically and without conscious awareness. After the provisional diagnosis is made, the radiologist undertakes a search for radiographic features that support this provisional diagnosis. This method is also called “pattern recognition”. Because of the experience of the radiologist, he or she can make comparisons of a case with previous cases already seen. A study by Brooks et al. has shown a clear association between diagnostic accuracy and familiarity with previous cases. There was an approximately 40% improvement in the diagnostic accuracy with the experienced group and approximately 28% with the novice group when these groups were looking at similar cases.

A study by Norman et al. tested two different approaches to interpretation of the electrocardiograph (ECG) with novices. Twenty-four undergraduate psychology students were given basic training on how to interpret an ECG. The students were divided in two groups: the forward reasoning group with sixteen subjects and the backward reasoning group with eight subjects. Participants in the forward reasoning group were asked to identify features from the ECG data and then use this list of features to generate a diagnosis while participants in backward reasoning group were asked to generate a diagnosis first and then identify features supporting their diagnosis. The diagnostic accuracy was 61.9% for the backward reasoning group and 49.4% for the forward reasoning group, and this difference was found to be statistically significant. The
forward reasoning group identified 5.22 features per ECG while the backward reasoning identified only 3.87. Moreover, the backward reasoning group identified fewer incorrect or irrelevant features. These differences were both found to be statistically significant. They concluded that using the forward reasoning method is detrimental compared to the backward reasoning.

Reviewing these studies, and more specifically, the Baghdady study, it is clear there are some limitations to all of them. The first limitation is the controlled nature of both the teaching and the presented cases; the teaching phase was all done one-on-one with each student using standardized computer software. In a regular classroom, the presence of more students and time constraint could impede the learning process. At the same time, the cases presented in the aforementioned studies were selected because they were classic examples of each abnormality. Obviously, a wide range of appearances is expected for a single abnormality which makes the diagnosis an even more complex process. Once again, the absence of patient information and clinical history could disadvantage the diagnostician. The next step in the evaluation of strategies in image interpretation consists of evaluating the actual search patterns used by the subjects. This will also confirm, or infirm, the use of a specific search algorithm by the dental students or other groups.

1.4 Eye Tracking

1.4.1 General Considerations

Eye tracking can be defined as "the process of watching where a person is looking". Eye tracking has been used since the 1960s in many fields, but most often in the publicity and medical domains. This process is used to determine how a subject looks at a specific stimulus, and then understand why they have looked there. Before eye tracking became available, researchers had to rely on subjects’ reports of their own eye movements. This clearly lacked objectivity and, depending on the study conducted, some subjects could describe their gaze in ways to satisfy the researchers rather than actually reporting their true gaze patterns. The advent of eye tracking systems changed this by allowing researchers a more objective and reliable method of monitoring the gaze of the subjects.
Contemporary eye tracking devices use a technique called pupil centre corneal reflection (PCCR). The systems focus an infrared beam on the cornea and pupil causing highly visible reflections. In turn, a camera captures the image of the eyes and the reflection patterns. Software then renders these numerical data into visual data. Multiple types of eye tracking devices have become available over the years. Some have been very large, often with a head stabilizer, providing very accurate tracking data. Some have been developed that are smaller and portable, allowing some liberty of movement from the observer, but with generally lower accuracy. Finally, some manufacturers have developed eye tracking glasses for studies in environments outside a laboratory setting (e.g. at the grocery store or while driving). Other than the size, the main differences between the systems have been their sampling or refresh rates (typically measured in Hertz), tracking range or distance, and more importantly their accuracy and precision of individual eye movements (both measured in degrees).

A number of unique descriptors have been used to describe the movements of the eyes detected by eye tracking systems. A fixation is defined as "the act or an instance of focusing the eyes upon an object"; how long does the eye remain on a single location. A saccade is defined as a "small rapid jerky movement of the eye especially as it jumps from fixation on one point to another", and represents path of eye movement made by the observer between each fixation (Figure 1). Blinking (or a blink) is "to close and then open your eyes very quickly". Blignaut and Wium also defined some more specific terms. They defined accuracy as "the distance between the actual and reported gaze positions", trackability as "the proportion of raw data samples that are lost during a recording", and finally precision as "the variance in position and latency measures".

Blignaut and Wium reported on the effect of ethnicity on the trackability, accuracy, and precision of eye movements. Their sample consisted of twenty-six people of African descent, twenty-two people of Southeast Asian descent, and twenty-three people of Caucasian descent. Subjects were shown a series of twenty-one tests, each consisting of a set of dots (similar to a calibration screen). For each individual test, the head position, the lighting condition, the gaze angle, or the stimulus background was modified. Their results showed that the trackability of the Asian group (87.2%) was lower than the African group (95.5%) and the Caucasian group (99.1%), being significant at $\alpha=0.1$ between the Asian and the Caucasian groups. As well, they
found that the precision was also worse for the Asian group (0.15°) than the other groups (both 0.11°), but only showed statistical significance at $\alpha=0.1$. As for the accuracy, they found it to be worse for the Asian group (0.91°) compared to the African group (0.57°) and the Caucasian group (0.61°); both showed statistical significance at $\alpha=0.05$. They attributed these differences to the narrower eyes of the Asian group. Another interesting finding was that all three groups performed worse in the dark, although this was only significant for the Caucasian group ($p=0.004$).

Multiple studies have been performed using eye tracking in the medical field; in the examination of an ECG, a radiograph, a histology slide or a clinical examination $^2, $ $^3, $ $^{12}$. Looking more specifically at the literature about radiology, eye tracking studies have found that radiologists do not see radiographs in the same way as novices (resident radiologists or students). Manning and his colleagues $^3$ recorded the differences between four observer groups with different levels of expertise looking for nodules on postero-anterior (PA) chest radiographs. The different groups were eight experienced radiologists, five experienced radiographers before and after six months of training in chest image interpretation, and finally eight novice undergraduate radiography students. The participants were asked to view 120 PA chest radiographs. They were then asked to decide whether lung nodules were present, and if so, to identify their locations. The authors then recorded the number of fixations, the length of the saccades, the visual coverage of the films and the mean scrutiny time per film. They conducted their experiment with a 504 remote optics system with magnetic head tracker® (Applied Science Labs, Bedford, MA, USA). Their results showed that the experienced radiologists and radiographers after training spent less time on each film, had less fixations with longer saccades and covered less area than the novice groups (the experienced radiographers before training and the undergraduate radiography students). They concluded that an increase of expertise comes with a greater efficiency and economy of effort. Matsumoto $^4$ compared the abilities of neurologists and other allied medical groups (nurses, medical technologists, psychologists and medical students), all with some knowledge about the brain but no training on brain imaging, to identify cerebrovascular accidents on brain computed tomographic (CT) examinations. Six brain CT images were shown to each observer, one of which was normal. The participants were asked to localize the lesion and to identify it, if possible. The authors recorded the time for participants to first look at the lesion and also the duration of fixation. The investigators also created heat maps to visualize their data in a more
qualitative way. The investigators found that while both groups looked at the high-attenuation (white) regions on the CT images, only the experts looked at the low-attenuation (black) regions. They concluded that the more experienced group ("experts") not only identified the region of pathologic abnormality, but additionally, they examined other structures they know could host disease. In contrast, less important structures (i.e. those structures the experts knew could not host disease) were passed over rapidly.

Krupinski ⁵ and Kundel et al. ⁶ compared the abilities of radiologists and radiology residents to detect calcifications or signs of cancer on mammograms. Like other publications of this genre, these investigators found that radiologists took nearly half the time compared to the residents to identify the abnormality, but as well, they also made some additional novel findings. The investigators noted that their participants required more time to make true positive and false positive diagnoses, and less time for a true negative diagnosis. This is easily explainable by the time taken by the subjects to think and analyze what they were seeing. In a similar vein, the investigators noticed that there were two different types of error: a recognition error (where the subject would fail to recognize the abnormality) and the decision error (where the subjects decided, after analysis, that they should not consider it abnormal). The total dwell time for the false positive and false negative cases that were associated with a recognition error were all shorter than the ones associated with a decision error. Both authors also identified an interesting analysis of what they call "global impression", which seems to be related to peripheral vision. They found that 67% of the identified cancers were fixated on within the first second, and that compared to the residents, the radiologists would rapidly fixate the second and third calcifications when they were present. Kundel et al. ⁶ stated that the "rapid fixation of potential cancer locations in the mammograms is best explained by a perceptual process in which a global analysis initiates and guides search". However, locating the abnormality is not sufficient to develop a correct diagnosis since only 63% of the initially fixated cancers were reported as such. In relation with these findings, Charness et al. ¹⁴ and Smith-Bindman et al. ¹⁵ developed the "deliberate practice" theory, whereby looking at hundreds of mammograms leads an individual to develop better perceptual mechanisms, and therefore a higher accuracy. Given that all of this work was specifically related to mammograms, the findings of similar studies in other fields ³, ¹⁴, ³⁴ have been similar, suggesting that "deliberate practice" in these other areas might lead to the same results.
In summary, these studies \(^4,^{11,23}\) highlight a number of reasons behind an experienced practitioner’s success in image interpretation; being able to picture-match an image with a large "mental encyclopedia" of known or similar-appearing entities, the ability to understand what is seen in a two-dimensional image and relate this to a three-dimensional rendering of a disease that reflects the "real" anatomy, physiology and pathophysiology of the patient \(^11\), and finally, the use of a holistic, non-analytical, backward reasoning model \(^4\).

### 1.4.2 Dentistry

In 2001, Suwa et al. \(^1\) published an eye tracking of CT images of the head and neck made for dental purposes. In this study, the investigators compared the abilities of dental specialists (oral and maxillofacial radiologists and oral and maxillofacial surgeons) to examine normal CT images of the head and neck region with CT images with pathology. The comparisons were made by analyzing six different parameters: total time to determine the absence or presence of pathosis, total number of fixations, total distance between the fixations, average time on each fixation, total fixation time, and the maximum fixation time. Their results showed a statistically significant difference between the normal images and those with pathology for three parameters: total fixation count, average time of the fixations, and total time. For all these, the subjects spent more time, and made longer and greater numbers of fixations on the normal images. There was a similar trend for the other three parameters, but the differences were not statistically significant. Their conclusion was that the differences between the normal and abnormal CT images could be explained on the basis that for a normal CT images the participants used a forward reasoning approach, compare to a backward reasoning method for CT images with pathology.

### 1.4.3 Data Output

Eye tracking data can be analyzed both qualitatively and quantitatively. Qualitative data is recorded by the eye tracking system as time to accomplish a certain task, duration of a gaze in a certain region of the image, number of fixations, saccades and blinks, and distance covered on the image. These variables give the investigator the ability to compare the performance of multiple observers performing a specific task. The qualitative analysis allows investigators to transform quantitative data into visual data, and then displays this over the images presented.
Depending on the type of analysis selected, this could represent the visualization of a scan path of eye fixations and duration or search pattern.

1.5 Panoramic Radiography

A panoramic radiograph is a tomographic image of the maxillofacial complex. To achieve this, an x-ray source and an image receptor rotate around the patient's head through a series of moving centre of rotation. The movement of the centre of rotation defines the focal trough through a curvilinear volume that includes the jaws. The structures within this focal trough are displayed clearly on the image receptor, while the structures outside the focal trough (i.e. in front of or behind) are blurred.

The panoramic radiograph displays the dentition and surrounding structures (temporomandibular joints, mandible, maxillae, and temporal bones) in a single image that permits the gross evaluation of these structures. As well, panoramic imaging utilizes relatively low radiation doses, approximately 20 μSv (between 15 and 25 μSv) compared to 171 μSv for a full mouth series (American National Standards Institute rated F-speed film with round collimation)\(^\text{31,32}\). On the other hand, panoramic radiography has inherently low image resolution (making it impossible to see fine details like early carious lesions), unequal magnification, and images artifacts. The last two factors, unequal magnification and image artifacts, can be kept to a minimum by correctly positioning the patient\(^\text{7,25}\).

The most common positioning errors are those that occur when the patient is positioned too far anteriorly or too far posteriorly within the imaging system. Patient positioning errors can lead to horizontal minification (anterior position) or magnification (posterior position). Another common error occurs if the patient’s head is tilted upwards or downwards. Upward tilting of the head can result in a flattened depiction of the occlusal plane, increased intercondylar distance, elimination of the posterior surface of the condylar head from the image, and superimposition of the hard palate over the roots of the maxillary teeth. When the head is tilted downwards, there may be excessive curvature of the occlusal plane, reduction in the intercondylar distance, elimination of the superior surface of the condylar heads and the inferior surface of the chin from the image, excessive tooth overlap, and hyoid bone superimposition over the mandible\(^\text{7}\).
Ghost images are inherent features of panoramic radiography and cannot be completely eliminated, even with perfect patient positioning. These images are created when the x-ray beam intercepts a structure before the center of rotation. Because of this, the image projected on the receptor will be positioned on the opposite side of the image in a position that is more superior to the actual anatomical structure, and more blurred than a structure within the focal trough.

Taken together, panoramic images are very complex radiographic images. It is therefore very important for dental students as well as practicing dentists to be familiar with, and be comfortable interpreting panoramic radiographs, especially since their use has been on the rise for the last 20 years. This familiarization with the panoramic images is also necessary to avoid false positive diagnoses. As was stated previously, Khalifa found an increase in false positive diagnosis when the subjects were trained to use a systematic search method. As the author points out, this is most likely because of a lack of comprehension of the way panoramic radiographs are acquired and therefore how the structures are represented on the radiographs. A basic understanding of both the physics of the image acquisition and of the anatomy is essential to exploit the full potential of panoramic radiography.

1.6 Summary

The medical radiology literature has demonstrated that experienced radiologists and students view radiographs differently. In general, experienced practitioners spend less time and have fewer eye movements than novices, and concentrate on the regions they know to be of interest more quickly than novices. These differences have been explained by the difference in the approaches these two groups use; backward and forward reasoning. In dentistry, even though only one paper has been published, the authors came to a similar conclusion. Suwa et al. also found a trend where the pathologic images were more often analyzed with a backward reasoning approach, while the normal images were analyzed with a forward reasoning approach. Unfortunately, they used static computed tomography (CT), which is not a common type of image used in dentistry.

1.7 Aim and Statement of Problem

The goal of this research is to analyze panoramic image search strategies in dental students and certified oral and maxillofacial radiologists. Panoramic radiographs are complicated tomographic
images containing many important anatomical structures. Their interpretation is complicated by the presence of both real and ghost images, overlapping anatomical structures, and images of soft tissue that can readily complicate image interpretation. Since these radiographs are used commonly in dentistry by general and specialist clinicians alike, it is vital that students learn to competently view, identify normal and abnormal features, and make interpretations that make sense so that patient care is optimized. Understanding the experienced practitioners’ method of image interpretation may enable us to develop better approaches in how we teach dental students.

1.8 Objectives

1. To analyze and compare the strategies that dental students and oral and maxillofacial radiologists use when viewing panoramic radiographs.

2. To determine if there is a common method that dental students and oral and maxillofacial radiologists use when searching a panoramic radiograph.

3. To identify image interpretation strategies that oral and maxillofacial radiologists use that can be applied to the education of dental students.

1.9 Hypotheses

1.9.1 Alternate Hypotheses

1. Certified oral and maxillofacial radiologists spend less time looking at the panoramic radiograph than dental students.

2. Certified oral and maxillofacial radiologists have fewer fixations on the panoramic radiograph than dental students.

3. Certified oral and maxillofacial radiologists cover less distance on the panoramic radiograph than dental students.

4. Certified oral and maxillofacial radiologists have fewer blinks on the panoramic radiograph than dental students.

5. Certified oral and maxillofacial radiologists have fewer and shorter saccades on the panoramic radiograph than dental students.
6. Certified oral and maxillofacial radiologists **identify an area of interest quicker** than dental students.

7. Certified oral and maxillofacial radiologists require **fewer eye fixations on an area of interest** than dental students.

8. Certified oral and maxillofacial radiologists **spend less time viewing an area of interest** than dental students.

9. Certified oral and maxillofacial radiologists have **revisits an area of interest** less often than dental students.

1.9.2 Null Hypotheses

1. There is no difference between certified oral and maxillofacial radiologists and dental students in regards to the **total duration of panoramic radiograph examination**.

2. There is no difference between certified oral and maxillofacial radiologists and dental students in regards to the **number of fixations on the panoramic radiograph**.

3. There is no difference between certified oral and maxillofacial radiologists and dental students in regards to the **distance covered on the panoramic radiograph**.

4. There is no difference between certified oral and maxillofacial radiologists and dental students in regards to the **number of blinks on the panoramic radiograph**.

5. There is no difference between certified oral and maxillofacial radiologists and dental students in regards to the **number and length of saccades on the panoramic radiograph**.

6. There is no difference between certified oral and maxillofacial radiologists and dental students in regards to the **time it takes to identify an area of interest**.

7. There is no difference between certified oral and maxillofacial radiologists and dental students in regards to the **number of eye fixations on an area of interest**.

8. There is no difference between certified oral and maxillofacial radiologists and dental students in regards to the **time spent viewing an area of interest**.

9. There is no difference between certified oral and maxillofacial radiologists and dental students in regards to the **number of revisits to an area of interest**.
Figure 1. Scan path of a dental student of a normal panoramic radiograph. Example of saccades (lines) and fixations (circles).
Chapter 2

2 Materials and Methods

2.1 Observer Selection

Research ethics approval for this study was obtained from the University of Toronto Health Sciences Research Ethics Board (Protocol Reference number 28709) (Appendix A).

Dental students were recruited from the fourth year (DDS2014) class from the Faculty of Dentistry of the University of Toronto. This group was selected because they had received training on panoramic image interpretation through the Dental Procedure Education System (DPES) module on panoramic radiography, authored by Dr. Susanne Perschbacher and Dr. Mindy Cash and produced with the help of the Media Services of the Faculty of Dentistry at the University of Toronto. They also had some experience reporting panoramic radiographs. Multiple emails were sent through the Faculty email system to advertise this study to the students (Appendix B). A small proportion of these students were foreign-trained dentists enrolled in the International Dentist Advanced Placement Program (IDAPP). The program begins with a 5 month period of study before they enter the third year DDS program. In this study they were considered as being dental students, although a differentiation was made during the test between them and the regular stream dental students in case comparison between the two groups was to be made.

The experienced practitioner group included certified oral and maxillofacial radiologists and who were either Diplomates of the American Board of Oral and Maxillofacial Radiology (ABOMR) or Fellows of the Royal College of Dentists of Canada (RCDC) in Oral and Maxillofacial Radiology. The testing part of the experiment with the experienced practitioners was done at the 2013 annual session of the American Academy of Oral and Maxillofacial Radiology (AAOMR) in Los Angeles. The recruitment began one week prior to the event by sending an email to the "Oradlist" bulletin board to which most of the AAOMR attendees subscribe to. Some attendees responded immediately with their intention to participate. During the event itself, the attendees were invited directly by the author and also by previous participants. There are currently one
hundred and twenty Diplomates of the ABOMR and eighteen RCDC Fellows; we had hoped to recruit approximately thirty participants (Appendix C).

2.1.1 Sample Size

As no similar studies had been performed previously in this area, there were no data available to calculate a sample size before the beginning of the study. By reviewing the medical studies that compared novices and experienced practitioners, most had ten or fewer participants per group, and statistically significant differences (p < 0.05) was commonly observed between groups.

Dreiseitl et al.\(^2\) used an eye tracking system to compare the search patterns of three different groups when assessing pigmented skin lesions. They showed 28 digital images of skin lesions to 16 subjects (dermatologists and dermatology residents) divided in three groups depending on their experience (less than two years, between two and five years, more than five years training in dermatoscopy). Nine participants were in the novice group, four in the intermediate group, and three in the experienced group. Their data were comprised of the diagnostic time, the gaze length, the number of fixations, and the fixation time. Even with such a small number of participants, they had a p-value of less than 0.01 in all four categories when determining statistical significance between groups with an analysis of variance (ANOVA). Further analysis showed a lack of statistical significance between the novice and intermediate groups, but a p-value of <0.001 for seven of the other eight comparisons made using a t-test.

As was stated previously, Manning et al.\(^3\) used four observer groups divided as follows: eight experienced radiologists, five experienced radiographers before and after six months of training in chest image interpretation, and finally eight novice undergraduate radiography students. For their statistical calculations, they grouped the radiologists and the radiographers after the training together as the experienced practitioners group, and the radiographers before the training with the undergraduate students as the novice group. The participants were asked to search 120 PA chest radiographs for lung nodules. The authors observed the number of fixations, the saccadic amplitude, the visual coverage and the mean time per film. Their results show statistical significance with a p-value under 0.05 for all four parameters.

Matsumoto and his colleagues\(^4\) observed how two groups of 12 participants each, a group of neurologists and a group of medical practitioners with no training on brain imaging (including nurses and medical students), searched brain CT images showing signs of cerebrovascular accidents. Only six brain CT images were shown, one of which was normal and the other five had different pathology on them. The authors
noted the time before the first fixation at the lesion and also the duration of the fixations in the region. Their results show a lack of difference between the two groups when a clear lesion was involved, but they also showed a statistical difference with the more subtle lesions. Lastly, a publication by Krupinski observed how three radiologists and three residents looked at mammograms, looking for calcifications or signs of cancer. Both groups were shown a set of 20 mammography cases, 15 had one or more microcalcification, while the other five were lesion-free. In terms of statistical data, the author only looked at the total dwell time on the images. The results showed a statistical difference (p < 0.01) between both groups for both types of images.

These publications give us some sample size guidance. But without any kind of previous data relating to this specific area of research, a power analysis could be reliably made. Therefore, after a pilot study of ten novice participants, a sample size calculation was performed. The calculations were done with the total dwell time on the radiograph, once per category (normal, obvious, intermediate, and subtle). While the mean and standard deviation for the novices could now be calculated, we could only estimate these for the experienced practitioners. Our calculations showed the need for approximately 30 to 40 subjects to attain statistical significance ($\alpha = 0.05$ and $\beta = 0.8$). This was considered to be the minimum needed to observe statistical significance between observer groups, but all willing participants were welcome and no upper limit for the number of participants was set. In the end, the sample size was 15 for the experienced practitioners group and 30 for the novices group was attained.

### 2.2 Selection of the Panoramic Radiographs

Twenty digital panoramic radiographs were selected for this study. Five of these images were normal and did not contain any pathoses. A panoramic radiograph was considered normal if it was free of bony pathoses, identifiable dental caries, severe periodontal disease or image production artifacts. The remaining fifteen panoramic images contained regions of pathosis. A panoramic radiograph was considered abnormal if it contained one or more pathoses. If a radiograph was determined to be abnormal, these images were classified further into three subcategories. These subcategories only related to the level of difficulty to detect the abnormality, not to diagnose it. This classification was based on a classification used by Khalifa. In it, three reviewers were asked to grade the subtlety of an abnormality on a panoramic radiograph from 1 to 10; 1 being "very subtle" and 10 being "very obvious". Khalifa does not
define “subtle”, but she does define “obvious” as an abnormality that “can be easily and quickly identified - without any searching”.

Over a hundred and twenty digital panoramic radiographs were selected from patient files in the Faculty of Dentistry at the University of Toronto and from the Department of Dental Oncology of the Princess Margaret Hospital, Toronto. These radiographs were anonymized and reviewed by an expert panel consisting of three certified oral and maxillofacial radiologists. These oral and maxillofacial radiologists were calibrated during a session where the main author and all three experts analyzed ten panoramic radiographs of different difficulty level. At the end of this session, all three experts agreed on the same level of difficulty for all ten images. On the basis of Khalifa’s work, three categories (subtle, intermediate, and obvious) were defined as follows:

**Obvious**: An obvious abnormality is one that is easily identifiable in the opinion of the expert panel. Furthermore, it is one that everyone (including dental students, general practitioners, and specialists) in the dental community should be able to identify.

**Intermediate**: An intermediate abnormality is one that may be missed by some dental students and one that a dentist or dental specialist should see with a bit of effort. This abnormality should be easily seen by an oral and maxillofacial radiologist.

**Subtle**: A subtle abnormality is something 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.

A panoramic radiograph was rejected if it was of insufficient quality, had identifiable dental caries, severe periodontal disease or image artifacts (Appendix D).

Each member of the expert panel individually reviewed the panoramic radiographs within a PowerPoint® (Microsoft Corp., Redmond, WA, USA) presentation on their personal computer. A conventional Delphi panel method was used where a 100% agreement between all three experts was required to accept a radiograph. After the first round of viewing, all the radiographs with 100% agreement were selected. Radiographs where all three experts disagreed were eliminated. The radiographs where a single expert disagreed with the other two were sent back to this expert along with the others' classification; if the expert changed his or her mind, the radiograph was
selected, otherwise it was eliminated. After a second round of viewing, a consensus was reached on 18 normal, 13 abnormal (obvious), 8 abnormal (intermediate), and 5 abnormal (subtle) panoramic radiographs.

2.3 Eye Tracking System

We used the RED-m® (Sensomotoric Instruments, Teltow, Germany) eye tracker system. This small eye tracker (24 cm by 2.5 cm by 3.3 cm, weighing 130 g) offers a sampling rate of 120 Hz, which is sufficient, since our subjects could maintain a steady head position during the experiment. No head stabilization devices were used since we wanted the subjects not to think about the eye tracker during the experiment. The operating distance between the device and an observer’s eyes is between 50 cm and 75 cm, with 60 cm to 65 cm being the best position. At 60 cm, the tracking area of the system is 32 cm by 21 cm, which allows for some head movement. 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, therefore 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 at the base of the screen of a 15.6-inch laptop (Dell Latitude E6530, Round Rock, TX, USA) with a magnetic strip adhered on the computer and a small magnet attached to the back of the eye tracker. This laptop was also used by subjects to answer the questions related to each image, and has a display resolution of 1600 by 900 pixels. A second screen (connected via VGA cable) was used by the principal investigator to ensure subjects stayed within the tracking range and operating distance. This second screen was a 19-inch (Dell 1909 Wf, Round Rock, TX, USA) with a display resolution of 1440 by 900 pixels.

The software associated with the eye tracker system is the Experiment Center® 3.3 (Sensomotoric Instruments, Teltow, Germany). This software can easily be used to build an experiment with multiple types of stimuli, including images. The software has the ability to randomize the 20 images for each participant the order of the stimuli and the associated questions (Figure 2). A nine point initial on-screen calibration was used for each participant, and this was followed by a four point calibration to confirm the preliminary calibration (Figure 3).
2.3.1 Software Programming

The first six screens of each experiment were used to obtain epidemiologic information about each participant (sex, age, educational background and years as a practicing oral and maxillofacial radiologist, if applicable). After these data were acquired, the next one hundred screens consisted of the twenty panoramic radiographs and their associated questions, randomized as a group. The questions for each of the radiographs were the following:

i. Is there an abnormality on this radiograph?

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

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

iv. On a scale from 1 to 10, rate your confidence level with regards to your interpretation/diagnosis of this radiograph.

Finally, the last screen was to let them know they were finished and to thank them.

2.4 Experimental Protocol

The experiment was conducted at the Faculty of Dentistry, University of Toronto and the Beverly Hills Hilton Hotel in Los Angeles, USA. Both venues were lit by fluorescent lights and could not be dimmed. The laptop's screen was placed approximately 60 to 65 centimeters from the participant, while the operator was seated at 90° to the subject so that the subject's gaze, head position and movement could be seen on a second monitor (Figures 4 and 5). Before positioning the subject, the protocol of the experiment was explained to them by the principal investigator. If a participant agreed, a consent form was signed (Appendix E). At that point, the principal investigator provided the subject with an explanation about the eye tracker, tips on how to ensure optimal gaze tracking (being centered within the field of the camera at a distance of approximately 60 cm to 65 cm) and how to operate the software. At this point, each participant was told to look at the images as if they were in their office, with a patient without any history or background information. They were told there was no time limit but, since they had a patient in the chair, they had to take a reasonable amount of time. Finally, they were told that some images may contain none, one, or more pathoses, without specifying the number of images for each.
Each participant was identified in the software with a code, either starting with a "P" for the novices or an "R" for the experienced practitioners, followed by two numbers. After an initial calibration, the software showed the accuracy of the gaze tracking and the operator could accept the calibration or reject it, and elect to re-calibrate the system (Figure 6). A study by Nodine et al. have determined that an accuracy of \( \leq 0.6^\circ \) is acceptable, while another study by Matsumoto et al. considered the calibration as successful if the maximum spatial error was < 1° and the average < 0.5°. For this experiment, a calibration result of < 1° was considered accurate. Once the experiment was completed, the operator gave feedback to the subject about missed abnormalities or interpretation, recurrent mistakes, or any other questions the subject might have had. The participants were also asked not to discuss the cases with other potential participants.

2.5 Statistical Analysis

A mixed effects analysis of variance (ANOVA) that included group (novice vs. experienced) as the between-subjects factor and type of radiograph (normal, obvious, intermediate or subtle) as the within-subjects factor was performed using SPSS® (IBM, Endicott, NY, USA). Given the number of comparisons in these analyses, we used the Bonferroni correction and thus set the alpha level at \( p < 0.005 \).
**Figure 2.** Screenshot of the Experiment Center showing the group randomization.
**Figure 3.** Points of the calibration screen.

**Figure 4.** Photograph of the setup. The participant is to the left, the observer is to the right.
**Figure 5.** Photograph of the setup (close-up).

**Figure 6.** Results of calibration.
Chapter 3

3 Results

The data for each participant were exported from BeGaze® (Sensomotoric Instruments, Teltow, Germany) to Excel® (Microsoft Corp., Redmond, WA, USA) where they were grouped according to participant type (novice or experienced practitioner), whether the radiograph was normal or abnormal, and level of difficulty (obvious, intermediate, subtle). The data were then exported into SPSS® (IBM, Endicott, NY, USA), and an ANOVA test was performed as described previously. 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. These two types of image were used to do descriptive analysis.

Please refer to Table 1 for a complete summary (means and standard deviation) of the results.

3.1 Quantitative Results

3.1.1 Total Time Examining the Radiograph

There was a significant main effect of type of radiographs, $F (3, 129) = 10.27, p < 0.001$. This effect tells us that if we ignore the level of expertise of the participants, there was a significant difference between some types of radiographs. The post-hoc test shows that there were significant differences between the normal images and all three others classification: obvious ($p < 0.001$), intermediate ($p = 0.002$), and subtle ($p = 0.03$). In all cases, the participants took longer to search the normal images than the abnormal images. Also, there was no significant effect of level of expertise, $F (1, 43) = 3.405, p = 0.072$.

Furthermore, there was a significant interaction between the classification of the images and the level of expertise of the subject, $F (3, 129) = 8.217, p < 0.001$. This effect tells us that there was a difference in the total time spent searching the images in regards to the classification of the images and the level of expertise of the participants. More specifically, post-hoc analysis using Tukey-Kramer approach showed there was a significant difference between the normal images and the others, solely for the OMRs, where they spent more time searching the normal images. There were also differences between the two groups in the three abnormal categories, where the
OMRs spent less time searching than the students. There was no difference between the two groups in the normal category.

### 3.1.2 Number of Fixations on the Radiograph

There was a significant main effect of type of radiographs, $F (3, 129) = 12.365, p < 0.001$. This effect tells us that if we ignore the level of expertise of the participants, there was a significant difference between some types of radiograph. The post-hoc test shows that there were significant differences between the normal images and all three others classification: obvious ($p < 0.001$), intermediate ($p < 0.001$), and subtle ($p = 0.02$). In all cases, the participants had more fixations in the normal images compare to the abnormal images. Also, there was no significant effect of level of expertise, $F (1, 43) = 2.423, p = 0.127$.

Furthermore, there was a significant interaction between the classification of the images and the level of expertise of the subject, $F (3, 129) = 8.975, p < 0.001$. Post-hoc analysis using Tukey-Kramer approach showed there was a significant difference between the normal images and the others, solely for the OMRs, where they had more fixations on the normal images. There were also differences between the two groups in the three abnormal categories, where the OMRs had fewer fixations than the students. There was no difference between the two groups in the normal category.

### 3.1.3 Distance Covered on the Radiograph

There was a significant main effect of type of radiographs, $F (3, 129) = 13.45, p < 0.001$. This effect tells us that if we ignore the level of expertise of the participants, there was a significant difference between some types of radiograph. The post-hoc test shows that there were significant differences between the normal images and all three others classification: obvious ($p < 0.001$), intermediate ($p < 0.001$), and subtle ($p = 0.06$). In all cases, the participants covered more distance in the normal images compare to the abnormal images. Also, there was no significant effect of level of expertise, $F (1, 43) = 0.151, p = 0.699$.

Furthermore, there was a significant interaction between the classification of the images and the level of expertise of the subject, $F (3, 129) = 14.362, p < 0.001$. Post-hoc analysis using Tukey-Kramer approach showed there was a significant difference between the normal images and the
others, solely for the OMRs, where they covered more distance on the normal images. There was also a difference between the two groups in the obvious category, where the OMRs covered less distance than the students. There was also a significant difference in between the two groups in the normal category, but in this case the OMRs covered more distance than the students.

### 3.1.4 Number of Blinks on the Radiograph

There was no significant effect of type of radiographs, $F(3, 129) = 0.516$, $p = 0.672$, and no significant effect of level of expertise, $F(1, 43) = 2.946$, $p = 0.093$. Furthermore, there was also no significant interaction between the classification of the images and the level of expertise of the subject, $F(3, 129) = 0.501$, $p = 0.683$.

### 3.1.5 Number of Saccades on the Radiograph

There was a significant main effect of type of radiographs, $F(3, 129) = 11.951$, $p < 0.001$. This effect tells us that if we ignore the level of expertise of the participants, there was a significant difference between some types of radiograph. The post-hoc test shows that there were significant differences between the normal images and all three others classification: obvious ($p < 0.001$), intermediate ($p = 0.001$), and subtle ($p = 0.019$). Once more, in all cases the participants had more saccades in the normal images than in the abnormal images. Also, there was no significant effect of level of expertise, $F(1, 43) = 2.912$, $p = 0.095$.

Furthermore, there was a significant interaction between the classification of the images and the level of expertise of the subject, $F(3, 129) = 8.255$, $p < 0.001$. Post-hoc analysis using Tukey-Kramer approach showed there was a significant difference between the normal images and the others, solely for the OMRs, where they had more saccades on the normal images. There were also differences between the two groups in the three abnormal categories, where the OMRs had fewer saccades than the students. There was no difference between the two groups in the normal category.

### 3.1.6 Length of Saccades on the Radiograph

There was no significant effect of type of radiographs, $F(3, 129) = 3.458$, $p = 0.018$, and no significant effect of level of expertise, $F(1, 43) = 4.049$, $p = 0.050$. Furthermore, there was no
significant interaction between the classification of the images and the level of expertise of the subject, $F(3, 129) = 3.973, p = 0.010$.

### 3.1.7 Time Before First Fixation in Area of Interest (AOI)

There was no significant effect of type of radiographs, $F(3, 86) = 3.906, p = 0.024$. There was a significant effect of level of expertise, $F(1, 43) = 9.137, p = 0.004$. This effect tells us that if we ignore the classification of the radiographs, there was a significant difference between the two groups (novice and experienced). More precisely, the experienced group took only $7.4 \pm 1.4$ seconds to localize the abnormality compared to $12.6 \pm 1.0$ seconds for the novice group. Finally, there was no significant interaction between the classification of the images and the level of expertise of the subject, $F(3, 86) = 0.518, p = 0.598$.

### 3.1.8 Number of Fixations in AOI

There was a significant main effect of type of radiographs, $F(3, 86) = 5.799, p = 0.004$. This effect tells us that if we ignore the level of expertise of the participants, there was a significant difference between some types of radiograph. The post-hoc test shows that there was a significant difference between the obvious and subtle images ($p = 0.01$), in which case the participants had fewer fixations on the AOI classified as obvious. Also, there was no significant effect of level of expertise, $F(1, 43) = 0.051, p = 0.822$.

Furthermore, there was a significant interaction between the classification of the images and the level of expertise of the subject, $F(3, 86) = 10.261, p < 0.001$. Post-hoc analysis using Tukey-Kramer approach showed there was a significant difference between the images with obvious pathoses and the others, solely for the OMRs, where they had fewer fixations in the AOI on the images with obvious pathoses. There was also a difference between the two groups in the obvious category, where the OMRs again had fewer fixations in the AOI than the students.

### 3.1.9 Total Time Looking an AOI

There was a significant main effect of type of radiographs, $F(3, 86) = 6.862, p = 0.002$. This effect tells us that if we ignore the level of expertise of the participants, there was a significant difference between some types of radiograph. The post-hoc test shows that there were significant differences between the obvious images and the intermediate ($p = 0.007$) and subtle images ($p =$
0.02. For both cases, the participants spent less time in the AOI of obvious pathoses. Also, there was no significant effect of level of expertise, $F (1, 43) = 0.004, p = 0.951$.

Furthermore, there was a significant interaction between the classification of the images and the level of expertise of the subject, $F (3, 86) = 9.625, p < 0.001$. Post-hoc analysis using Tukey-Kramer approach showed there was a significant difference between the images classified as obvious and the others, solely for the OMRs, where they spent less time in the AOI on the images with obvious pathoses. There was also a difference between the two groups in the obvious category, where the OMRs again spent less time in the AOI than the students.

### 3.1.10 Number of Revisits in AOI

There was no significant effect of type of radiographs, $F (3, 86) = 1.758, p = 0.179$, and no significant effect of level of expertise, $F (1, 43) = 0.575, p = 0.452$. Furthermore, there was a significant interaction between the classification of the images and the level of expertise of the subject, $F (3, 86) = 5.592, p = 0.005$. Post-hoc analysis using Tukey-Kramer approach showed that the only significant difference was between the two groups in the obvious category, where the OMRs had fewer revisits in the AOI than the students.

### 3.2 Descriptive Results

#### 3.2.1 Heat Maps

Our results showed that students have a greater tendency to proportionally spend more time searching the regions of ghost images and superimpositions (like the cervical spine or the airway, for example) than the remainder of the image compared to OMRs (Figure 7 and 8). They also spent more time examining the dentition than the OMRs (Figures 9 and 10). These observations are true in all categories, but more so in the normal category. Normal variations of the anatomy (like the exostoses seen at the angles of the mandible caused by the insertion of the masseter muscles) are also more frequently examined by the students than the OMRs (Figures 11 and 12). Although the students paid more attention to these regions, they also spent less time looking at regions like the temporomandibular joints (TMJ) (Figures 7 and 8). In the case of larger lesions students did not cover as much of the lesion as the OMRs did (Figures 11 and 12). Another instance where the students did not cover as much of the image was in images where there were multiple abnormalities. Students looked at the abnormalities, but they did not search the
remainder of the image as thoroughly as they did on the other images to find other similar abnormalities (Figure 13). Interestingly, students compared the contralateral side to the affected side when an abnormality was detected; this was not done by the OMRs.

3.2.2 Scan Paths

When comparing the scan paths of a single student with him- or herself over multiple radiographs, it is clear that students do not follow a similar and consistent search pattern on each radiograph. Some do have a similar pattern on all the normal radiographs, but as soon as there is an abnormality, the search pattern changes completely and some regions that were included in the normal search pattern, are excluded.
Table 1. Summary of the quantitative results (mean ± SD).

<table>
<thead>
<tr>
<th>Parameters Evaluated</th>
<th>Radiographs Classification</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>DS</td>
</tr>
<tr>
<td>Total Time (s)</td>
<td>Normal</td>
<td>64.3 ± 5.5</td>
</tr>
<tr>
<td></td>
<td>Obvious</td>
<td>63.3 ± 4.8</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>62.7 ± 4.6</td>
</tr>
<tr>
<td></td>
<td>Subtle</td>
<td>61.7 ± 4.9</td>
</tr>
<tr>
<td>Number of Fixations</td>
<td>Normal</td>
<td>178 ± 16</td>
</tr>
<tr>
<td></td>
<td>Obvious</td>
<td>173 ± 14</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>171 ± 13</td>
</tr>
<tr>
<td></td>
<td>Subtle</td>
<td>173 ± 14</td>
</tr>
<tr>
<td>Distance Covered (cm)</td>
<td>Normal</td>
<td>702 ± 73</td>
</tr>
<tr>
<td></td>
<td>Obvious</td>
<td>709 ± 61</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>704 ± 57</td>
</tr>
<tr>
<td></td>
<td>Subtle</td>
<td>708 ± 59</td>
</tr>
<tr>
<td>Number of Blinks</td>
<td>Normal</td>
<td>20 ± 3</td>
</tr>
<tr>
<td></td>
<td>Obvious</td>
<td>23 ± 3</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>44 ± 19</td>
</tr>
<tr>
<td></td>
<td>Subtle</td>
<td>20 ± 3</td>
</tr>
<tr>
<td>Number of Saccades</td>
<td>Normal</td>
<td>187 ± 16</td>
</tr>
<tr>
<td></td>
<td>Obvious</td>
<td>181 ± 14</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>180 ± 14</td>
</tr>
<tr>
<td></td>
<td>Subtle</td>
<td>181 ± 14</td>
</tr>
<tr>
<td>Length of Saccades (ms)</td>
<td>Normal</td>
<td>26 ± 2</td>
</tr>
<tr>
<td></td>
<td>Obvious</td>
<td>26 ± 1</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>26 ± 2</td>
</tr>
<tr>
<td></td>
<td>Subtle</td>
<td>25 ± 1</td>
</tr>
<tr>
<td>Time before 1st Fixation (s)</td>
<td>Obvious</td>
<td>11.6 ± 1.2</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>15.4 ± 1.7</td>
</tr>
<tr>
<td></td>
<td>Subtle</td>
<td>10.9 ± 0.8</td>
</tr>
<tr>
<td>Number of Fixations in AOI</td>
<td>Obvious</td>
<td>15 ± 1</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>13 ± 2</td>
</tr>
<tr>
<td></td>
<td>Subtle</td>
<td>15 ± 2</td>
</tr>
<tr>
<td>Total Time in AOI (s)</td>
<td>Obvious</td>
<td>5.7 ± 0.5</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>5.4 ± 0.8</td>
</tr>
<tr>
<td></td>
<td>Subtle</td>
<td>5.5 ± 0.7</td>
</tr>
<tr>
<td>Number of Revisits</td>
<td>Obvious</td>
<td>8 ± 1</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>7 ± 1</td>
</tr>
<tr>
<td></td>
<td>Subtle</td>
<td>7 ± 1</td>
</tr>
</tbody>
</table>
**Figure 7.** Heat map of the dental students of an abnormal (intermediate) panoramic radiograph. There is a supernumerary tooth distal to the maxillary left third molar (2.8).

**Figure 8.** Heat map of the OMRs of an abnormal (intermediate) panoramic radiograph. There is a supernumerary tooth distal to the maxillary left third molar (2.8).
**Figure 9.** Heat map of the dental students of a normal panoramic radiograph.

**Figure 10.** Heat map of the OMRs of a normal panoramic radiograph.
Figure 11. Heat map of the dental students of an abnormal (intermediate) panoramic radiograph. There is a benign tumour (hemangioma) located distal to the maxillary right second molar (1.7).

Figure 12. Heat map of the OMRs of an abnormal (intermediate) panoramic radiograph. There is a benign tumour (hemangioma) located distal to the maxillary right second molar (1.7).
Figure 13. Heat map of the dental students of an abnormal (obvious) panoramic radiograph with multiple abnormalities. There are four keratocystic odontogenic tumours (KOTs) located in the following regions: maxillary left molar, mandibular left molar, mandibular left first premolar, mandibular right canine.
Chapter 4

4 Discussion

4.1 Evaluation of the Quantitative Results

4.1.1 Parameters Involving the Radiograph

An examination of the results from the total time spent searching the radiograph, the number of fixations, and the number of saccades, it is clear that there is a similar pattern of findings for these parameters. There were significant group differences noted for the three abnormal categories (obvious, intermediate and subtle). These differences showed that the OMRs spent less time, had fewer fixations, and had fewer saccades. These findings are consistent with our hypotheses. These data clearly show that OMRs are faster to visualize the abnormality. This is consistent with the study by Dreiseitl et al. \(^2\), where they found that experienced dermatologists could identify a pigmented lesion faster than those with less experience and training. The conclusion we draw from our findings is that like Dreiseitl et al., the increased level of experience of the clinician (i.e. OMRs) allows them to both find the abnormality and evaluate it faster. Supporting this interpretation, Manning et al. \(^3\) also showed that novice participants spent more time, had more fixations, and used more but shorter saccades. Manning et al. related these findings to the ability of the experienced practitioners to group larger regions of an image together. In contrast, novices may use a point-by-point examination method of the whole image.

On the other hand, there was no statistically significant difference between dental students and OMRs for the normal images in terms of total time spent searching the radiograph, number of fixations, and number of saccades. Yet, the analyses did show differences between the normal images and the others for the OMRS with the following parameters: total time spent searching, number of fixations, distance covered, and number of saccades. These differences could be explained by the OMRs performing a more thorough search of normal images compared to abnormal images, in order to completely exclude the possibility of an abnormality. For the distance covered, there are statistically significant differences in the normal and obvious categories of images. In the obvious category, the OMRs’ eyes traveled less distance than the students’. This again demonstrates the efficiency of the OMRs compared to the students. What is interesting is the lack of difference in the intermediate and subtle categories. This is most likely
because they had to do a more thorough search to first localize the abnormality, and then look around more to identify its features to interpret it correctly. On the other hand, for the normal category the OMRs covered 200 cm more than the students. This is interesting because there was no difference in this category for all the other parameters. So even though they did not spend more time, had more fixations, saccades, and blinks, or longer saccades, they still managed to cover nearly 30% more than the dental students did.

4.1.2 Parameters Involving the Area of Interest (AOI)

With respect to the interpretation of abnormalities, there is only a difference in the obvious category between the OMRs and students. This is demonstrated by the fewer fixations, fewer revisits, and less time spent in the AOI by the OMRs. Since the diagnosis of the obvious category is easier for the OMRs, these significant differences are to be expected. An interesting finding is the trend seen for the intermediate and subtle categories, where the results are the opposite. Students spent less time, had fewer fixations, and fewer revisits on these AOIs than the OMRs, although these differences are not statistically significant. Another interesting finding is the similar difference in the number of fixations and time spent in the AOI by the OMRs between images in the obvious category compared to intermediate and subtle pathoses. In both cases, OMRs spent less time and had fewer fixations with obvious pathoses. This again relate to the ease of diagnosing obvious pathologies for OMRs. Finally, there was a difference overall between the two groups in the time before the first fixation in the AOI. This shows not only the ability of the OMRs to identify an abnormality faster, but demonstrates that they focus their gaze on the regions they know to harbor pathoses. This finding is consistent with the conclusion by Matsumoto and his colleagues, who noted that experts (neurologists in their case) searched more inconspicuous, but clinically relevant regions, than their novice counterparts. This suggests that the speed of the OMRs is due, at least in part, to their knowledge and experience. These two factors led them to search areas on the radiographs where they felt pathoses were more likely to occur. In contrast, OMRs might glance more quickly at other areas where they know the probability of disease is lower or nil.

4.2 Evaluation of the Descriptive Results

Many of the findings relating to the portrayal of the data as heat maps can be related to the lack of understanding or experience of the students. For example, the heat maps show that the
students fixate their gaze in the regions of the ghost images and anatomical superimpositions; technical artifacts of image production. The main region where this effect was seen is the ghost image of the cervical spine, in the centre of the radiographs. This may reflect a lack of understanding of the physics behind the acquisition of this type of radiograph or a lack of appreciation for how normal anatomy appears, both inside and outside of the focal trough. Another major difference between the students and the OMRs is the absence of fixations in some regions, and an overabundance of fixations in some other regions of the images by the students. More specifically, the students had a tendency to overlook some regions like the TMJs, the soft tissues of the neck and the cervical spine, but fixated on the dentition which may be very familiar to them. Although this is not too surprising, it is in a sense disappointing since they all received the same training in panoramic image interpretation in which they were taught to search and examine every aspect of the panoramic image in a systematic manner.

Similar comments could also be made about the scan paths. Students’ scan paths varied from radiograph to radiograph; it was inconsistent. Since the objective of a systematic search is to ensure that every region of the image will still be searched, regardless of the potential to identify an abnormality, it would be expected that the scan path of a single subject over multiple radiographs would be similar. This was found not to be the case, especially when the abnormality was obvious, and even more so when there were multiple obvious abnormalities. When comparing the standard scan paths of the students to the scan paths they used on a particular radiograph, it is obvious that their scan paths were severely altered. This is most likely due to something called the satisfaction of search. This term is predominantly used to describe the failure of an observer to detect a second lesion on the same radiograph after the first abnormality has been found. This definition can also be expanded to include the absence of search after the detection of the first lesion. To counter that, the students are also taught not to analyze an abnormality as soon as they see it; instead, they should continue with their normal search path to cover the remaining of the image and then come back to the abnormality and then evaluate it. Unfortunately, the excitation that comes with the detection of an abnormality has a propensity to cause the students to discontinue their search, and miss any remaining structures on the radiograph. The importance of a thorough search pattern in the education of the dental students should be reinforced. This hopefully, may lead to a decrease in the "satisfaction of search" effect, and at the same time ensure that all structures are thoroughly searched. Although
this does not mean that all the abnormalities on a radiograph would be detected since a deeper understanding of the appearance of normal versus abnormal is required. One way the students have developed to compensate for this deficiency is to compare the contralateral side when they suspect an abnormality. This can be seen in multiple students' scan paths and also in some, but few, OMRs' scan paths. This finding is consistent with a conclusion by Krupinski where she noted that readers with experience would engage in fewer comparison scans than those with less experience, at a ratio of 1:3.

### 4.3 Study Limitations

Although this study provides us with some insight and a better understanding of how OMRs and dental students search panoramic radiographs, the image viewing conditions did not exactly simulate the conditions that may occur in a dental office. One of these differences may be the lighting condition of the room; as it was stated previously it was not possible to completely lower the lights in the rooms used for this experiment. This could have made the detection of small or low contrast abnormalities more difficult. Another confounder was that all the images were shown using software in which no manipulation of the contrast and brightness was possible. As well, participants could not magnify areas within the image. To compensate for this, all selected images were acquired digitally to avoid loss of information in the digitization process that can occur with film-based imaging. The images were all pre-processed by the principal investigator to adjust the contrast, brightness, and to compensate for these if the image was over- or under-exposed. The incorporation of image viewing software with the eye tracking software could solve this problem. Another difference with a normal clinical setting was that the participants were not given any patient clinical history. Giving a medical history to the participants would make each image scenario more clinically relevant, but at the same time it could potentially bias the search pattern and the approach used by the subjects. Hatala et al. demonstrated this effect in another visually-based domain; interpretation of the electrocardiogram (ECG). They showed ECGs with some information about the patient, like age, sex, and profession. They demonstrated that the subjects were influenced by this information and would diagnose the ECGs the same way if this information were similar between two patients.

Interestingly, the subjects were expecting abnormalities on the radiographs. Even though they were told a certain number of images would be free of abnormality (i.e. “normal”), they may
have felt the need to search more thoroughly since they were under scrutiny. Because of this, it is not unreasonable to assume that both groups took longer than they normally would in their normal clinical setting. Also, some students may have forced themselves, for the same reason, to perform the systematic search that had been taught to them. Unfortunately, even if this experiment was repeated in a clinical environment, the subjects would still know they are being observed and the same bias may occur. It would therefore be impossible to perform this experiment differently, partly because of the limitations of the eye tracker and the knowledge of its presence by the subjects.

Another fact that may affect the results is the degree to which each subject can appreciate areas peripheral to a point of fixation registered by the system. It is possible for a subject to visualize areas just outside of an AOI but still be able to search within the AOI with their peripheral vision. This was previously demonstrated by Matsumoto and his colleagues \(^4\) when they noted that subjects practically never looked at the edge of the cranium, but they were still able to identify it. They found this to be true specifically for areas of greatest contrast difference in an image.

Finally, the ethnicities of the participants were not taken into account. It was found by Blignaut and Wium \(^18\) that the trackability, accuracy, and precision for Asian participants were worse than African and Caucasian subjects. Even though there is no way to control the ethnicity of the participants (ethnicity should not be a limiting factor in the selection of the participants), more care should be taken when positioning and calibrating Asian participants to get the best possible accuracy, precision, and trackability.

### 4.4 Implications for Oral Radiology

There are two equally important components to radiology: localization and interpretation. Localization relies on a deep knowledge of the normal anatomy and its variants, and also on participants using a solid search algorithm. As we demonstrated, the students have only partially acquired these skills. Fortunately, these skills can and will likely be acquired throughout their careers. Of course everyone requires a basic level of knowledge in order to be able to create a mental representation of the range of normal; students need to interpret many more radiographs than they do during their training. But this is not something specific to radiology; it is actually a
common observation in all disciplines of dentistry. This idea is supported by the observed differences between the RS students and the IDAPP students. The IDAPP students, with more experience than the RS students, were faster and more efficient when searching the radiographs. It would therefore make sense to expose the students to as many radiographs as possible to help them build a mental encyclopedia. In doing so, it would be important to emphasize the significance of a standardized search pattern to assure a complete and thorough search of all the structures visible on the radiograph, at the same time giving the students knowledge about the wide variety of normal anatomy. A specific area that this study has shown to be a weak for the students is the depiction of ghost images and superimpositions of overlying structures. Once more, the interpretation of more radiographs could lead to an understanding of the range of what is to be expected from these projections on the resulting image. This should be done under supervision to allow for direct interaction and to accelerate the learning process.

As for the second part, the interpretation, experience alone is not enough. Not only the knowledge of the radiographic features of the pathoses is necessary, but as Baghdady recently showed, a deep understanding of the basic sciences underlying the etiopathogenesis of a pathosis' appearance increases the diagnostic accuracy. This means that, alongside the normal material being taught in oral radiology classes, the pathophysiology of the diseases should be integrated in this teaching so as to underpin the reason for why a particular feature is seen. This was also the conclusion of Manning et al. when observing radiologists search lung nodules on PA chest radiographs. They conclude that it is not only the deeper knowledge of the experienced practitioners, but how they use it that led to their better performance. They emphasize the need for a deeper understanding of the pathophysiology instead of exposing the students to a large number of cases.

These findings could be applied to the way radiographic interpretation (more specifically regarding panoramic radiography) is taught. The main features of the OMRs' results show that they don't "waste" their time looking at regions of ghost images or superimposition, they know the range of normal, and most of them apply a systematic search of the images. Obviously, some of these differences are more likely to be explained by the difference in the level of expertise of these two groups and therefore cannot be taught to the students. The use of a systematic search is something that, even though it is taught at the moment, could have a greater emphasis put on.
shortened and condensed video of what is already shown to the students could be created and the students would have to watch it before each panoramic interpretation. Watching this video multiple times makes it more likely for the students to follow a systematic search strategy every time and also create a subconscious habit for them to always do it the same way.

4.5 Future Directions

This study gives us insight in how dental students and oral and maxillofacial radiologists search panoramic radiographs. As we have demonstrated, there are multiple differences in the way these two groups undertake this task, and some changes are suggested to narrow these differences and improve the diagnostic ability of the students. Khalifa's study \(^ {16}\), where she compared the diagnostic efficiency of a systematic search strategy, demonstrated an increase in the true-positive (but also false-positive) diagnoses. As was suggested in her research, a combination of her work and the present study would allow us to evaluate the efficacy of this systematic search strategy from another point of view and at the same time confirm the use of the search strategy.

Another study that could benefit from the use of an eye tracker is the work by Baghdady \(^ {23}\) where she asked two groups of students, one defined as “analytic” and the other as “non-analytic”, to detect features on radiographs. Her preliminary results show the students in the analytic group reporting less irrelevant features. This could mean they have a more focused search strategy than the other group. Obviously, a combination of what was used in this study with what she did could show without a doubt the actual search strategy used by the two groups. This, along with the number of relevant or irrelevant features reported could shed even more light on the difference between these two strategies. For example, one could expect the participants using the backward reasoning method to spend less time, fewer fixations, and fewer revisits in AOI because of a picture matching process rather than a step by step analysis.

The main finding of this study was that experience level was responsible for the differences shown between the students and the OMRs. It would be interesting to repeat a similar study with more varied groups, including general dentists with a range in the number of years of practice. This could allow us to find how many years are needed to become an "expert" in radiology.
In the future, these studies could use more specific parameters than the whole range we looked at. By looking at our results, we can appreciate that the duration of the saccades and the number of blinks were not useful parameters for identifying differing levels of experience. As for the remainder of the parameters looking at the images, they only showed differences between the normal images and the three abnormal categories. That is, these parameters did not differentiate between the different classifications of pathoses. In terms of the parameters looking at the AOI, the time before the first fixation and the number of revisits lacked significance between the classifications of images, although the time before the first fixation yielded the only significant difference between the two groups of participants for all parameters. On the other hand, there were some differences between the obvious pathoses and either the subtle pathoses (number of fixations) or both of the remaining category (total time spent in AOI). Depending on the type and objectives of future study, the authors should use parameters that will help them achieve these objectives. We recommend that a focus on eye-tracking parameters related to the area of interest may be a more fruitful approach, as these appeared to be the most sensitive to our manipulations in the present study design.

Since this was the first experiment of its kind in dentistry, we asked the participants a multitude of questions that, in the end, were not used in the analysis. First, it would be interesting to see if there are differences between male and female participants, since no previous study has looked at this. This would require a larger number of participants with a similar level of experience. Even though the emphasis was not on the diagnostic ability of the participants, they were asked for a diagnosis for each radiograph they identified as being abnormal. The relationship between the diagnostic accuracy and the search pattern could be used to confirm Khalifa’s finding that the use of a systematic search strategy leads to an increase in true-positive and false-positive diagnoses. This could also help determine if the participants mostly do search errors, recognition errors, or decision errors, as discussed earlier. Lastly, we also asked how confident they were for each interpretation. There may be a link between the level of confidence and the search strategy used by the participants. Fortunately, these questions could easily be answered in a subsequent study using the data collected during this one.

This study employed a commonly used radiograph, but as cone-beam computed tomography (CBCT) machines become more common, it would be interesting to evaluate the search pattern
used by the experienced practitioners. The information gathered from the experienced practitioners could be used to create a standardized search algorithm, similar to the one presented by White and Pharoah for the panoramic radiograph. This search pattern could then be taught in the omnipresent continuing education courses regarding CBCT. This could ensure a higher standard of care for the patients and a more accurate interpretation of the CBCT volumes.

A similar study could observe how participants search the features of a specific subset of pathoses, for example how they search different bone dysplasia. These results could show the features of each of the individual pathoses that draw the attention of the participants. Going even further, the same thing could be done with a single pathosis, for example an ameloblastoma. By showing multiple ameloblastoma, the results would demonstrate what specific features of the ameloblastoma attract the attention of the participants.

Finally, this study was set up in such a way that during the experiment the subjects were not provided assistance during the procedure. The eye tracker would be a great tool to give access into the minds of the students, and then provide instant feedback. This would definitely be impractical in most dental schools, simply because the sheer number of students wouldn't allow it. On the other hand, it could definitely be used as a teaching tool in graduate studies, like oral and maxillofacial radiology or even oral medicine and pathology, where the ratio between professors and students is much lower.
Chapter 5

5 Conclusion

This study set out to determine if there were differences in the ways in which dental students and oral and maxillofacial radiologists search panoramic radiographs. The results demonstrate that there is a significant difference in the efficiency of localizing pathoses between these two groups. Overall, the OMRs search the radiograph and the area of interest more quickly but still more thoroughly than dental students. This difference could be reduced by emphasizing the use of a systematic strategy by students. The strategy taught for the moment is adequate, but unfortunately it is not always utilized by the students. One recommendation may be to create a short video (2-3 minutes) that could be viewed by students before every panoramic radiograph interpretation. This might create a habit and possibly increase the number of abnormalities localized. The results also show that one of the main reasons for this is the difference in experience levels between the two groups as evidenced by the increased number of fixations on unimportant or irrelevant regions of the radiograph by the dental students. A repeat of this study with general dentists with a wide range in the number of years of practice could give us a better idea on what is necessary to become an “expert”. Future studies using an eye tracker should focus on specific parameters, depending on the goal of the study. For example, parameters like the total time, the number of saccades, the number of fixations and the total distance covered on the image could be used effectively to distinguish between normal and abnormal images. On the other hand, parameters like the number of fixations and total time spent in AOI could be used to distinguish between observers’ abilities to distinguish between different degrees of subtlety. Lastly, the time before the first fixation in the AOI could be used to distinguish between the expertise of different observer groups.
References


Appendix A

Health Sciences Research Ethics Board Approval Letter

UNIVERSITY OF
TORONTO

OFFICE OF THE VICE PRESIDENT, RESEARCH

PROTOCOL REFERENCE #: 28709

March 25, 2013

Dr. Ernest Lam
FACULTY OF DENTISTRY

Dr. Daniel Turgeon
FACULTY OF DENTISTRY

Dear Dr. Lam and Dr. Daniel Turgeon,

Re: Your research protocol entitled, "The influence of experience and training in the examination pattern of panoramic radiographs: An eye-tracking study"

ETHICS APPROVAL

Original Approval Date: March 25, 2013
Expiry Date: March 24, 2014
Continuing Review Level: 1

We are writing to advise you that the Health Sciences Research Ethics Board (REB) has granted approval to the above-named research protocol under the REB’s delegated review process. Your protocol has been approved for a period of one year and ongoing research under this protocol must be renewed prior to the expiry date.

Any changes to the approved protocol or consent materials must be reviewed and approved through the amendment process prior to its implementation. Any adverse or unanticipated events in the research should be reported to the Office of Research Ethics as soon as possible.

Please ensure that you submit an Annual Renewal Form or a Study Completion Report 15 to 30 days prior to the expiry date of your current ethics approval. Note that annual renewals for studies cannot be accepted more than 30 days prior to the date of expiry.

If your research is funded by a third party, please contact the assigned Research Funding Officer in Research Services to ensure that your funds are released.

Best wishes for the successful completion of your research.

Yours sincerely,

Judith Friedland, Ph.D.
REB Chair

Daniel Gyewu
REB Manager

OFFICE OF RESEARCH ETHICS
McMaster Building, 12 Queen's Park Crescent West, 2nd Floor, Toronto, ON M5S 1S8 Canada
Tel: +1 416 946-3273 · Fax: +1 416 946-5763 · ethics.review@utoronto.ca · http://www.research.utoronto.ca/for-researchers-administrators/ethics
Hello DDS2014!

I am a (soon-to-be) 3rd year resident in the Oral and Maxillofacial Radiology post-graduate program right here at UofT. I am starting my data collection for my master's project and I need your help to do so! Why your group you ask? Well first of all, since I had the chance to teach some of you in the radiology clinic this past year I know you're a great group! Also, it's because I will ask you to look at panoramic radiographs, and since the future third years haven't had a chance to look at panoramic radiographs and haven't had the panoramic seminar either, it would be too difficult for them.

Without giving you any specifics of what I will be looking at, I can still give you an idea of what is expected from you. You will be presented 20 digital panoramic radiographs and you will be asked to assert if there are any abnormalities or not on them, to localize these abnormalities, and, if possible, to give a diagnosis (but that's really not the important part, so don't worry!). Evidently, the time needed for this experiment will vary, but you can expect it to take around 30-45 minutes.

Unfortunately, the Research Ethics Board has denied our suggestion of "giving" those of you who would participate an extra 5% in your radiology class. Therefore, your participation is to be completely voluntary... although you will definitely get my outmost gratitude!

I am presently in the final step of building the experiment and I should be able to start on June 17. I will be here from June 17 to 21, and most of July and August (even though I assume most of you won't!) and I can easily accommodate you during the day or evening, except Tuesdays PM and Thursdays AM. I would very much like to see as many of you as I can during this summer, so I'll have some data to present on my poster in October! I know this is short notice, but the selection of the radiographs turned out to be more complicated than I thought it would.
If you would like to participate, I invite you to email me with your preferred availabilities and I will get back to you as soon as possible.

Thank you and I hope to see you soon!

Daniel Turgeon, DMD

Oral Radiology resident
Appendix C

Email sent to OMRS through “oradlist” for participation in the study

Dear American Diplomates and Canadian Fellows,

I am a third year resident at the University of Toronto and as part of my M. Sc. project, I am trying to compare how dental students and Oral and Maxillofacial Radiologists search panoramic radiographs. To do so, I am gathering data with an eye-tracker, while presenting 20 digital panoramic radiographs. The only way for me to gather a large enough sample is to ask you, the attendees of the AAOMR annual meeting, to participate. As a way to control the variability of the results, my committee and myself have decided to limit the participation to Oral and Maxillofacial Radiologists that are either ABOMR Diplomate or RCDC Fellow.

The experiment itself is expected to take between 15 to 30 minutes, and will vary depending on how much time you spend searching the radiographs. For those who are looking at images in a more instinctive way, this would be a great opportunity to learn exactly how you are doing it!

I will be located within the exhibit hall, and will be there from Wednesday to Saturday, during conferences and between them, to perform this experiment. If you are interested, you can email me at daniel.turgeon@dentistry.utoronto.ca and tell me when you would like to participate. Otherwise, I will be waiting for you in the exhibit hall.

For those you also interested in the results, I will be presenting the results of the pilot study I did this summer at the first poster session on Saturday.

Thank you, and see you in sunny Los Angeles!

Daniel Turgeon

PS: As a little thank you, I’ll have Coffee Crisp (Canadian chocolate bar) for the participants!
Appendix D

Instructions and example for the panoramic radiographs selection

First of all I would like to thank you for taking time to help me build this experiment. I hope that by the end of this project, we will know how clinicians, whether they are dental students, dentists, residents or specialists, search a panoramic radiograph.

To help me do so, I ask that you classify these following panoramic radiographs in different categories.

1. Normal: A panoramic radiograph is considered normal if it is free of bony pathoses, identifiable dental caries, major periodontal disease or image artifacts.

2. Abnormal: A panoramic radiograph is considered abnormal if there is a dental anomaly, an inflammatory lesion, a cyst, a tumour, or an osseous dysplasia present. Please do not consider dental caries, major periodontal disease or artifacts as abnormal. Following the decision to classify a radiograph as abnormal, you will have to classify this abnormality further between three subcategories. These categories only relate to the difficulty of detecting the abnormality, not diagnosing it.

   2a. Obvious: An obvious abnormality is something that everyone in the dental community should be able to observe without too much efforts.

   2b. Intermediate: An intermediate abnormality is something that a dentist with a minimum of experience should see with a bit of effort. It might be missed by some dental students, but should be easily seen by radiologists.

   2c. Subtle: An obvious abnormality is something that dental students and some dentists would overlook, and that even a radiologist needs to pay attention to notice it.

3. Rejected: A panoramic radiograph will be rejected if it is of insufficient quality, has identifiable dental caries, major periodontal disease or image artifacts.

If there is one (or more), please localize it by putting an X in the appropriate section(s), using the panoramic radiograph diagram.
Case # Example

<table>
<thead>
<tr>
<th>Normal</th>
<th>Abnormal</th>
<th>Localization</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Subtle</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td></td>
</tr>
<tr>
<td>Rejected</td>
<td>Obvious</td>
<td>X</td>
</tr>
</tbody>
</table>

**Figure 14.** Diagram of a panoramic radiograph for localization of abnormality.
Appendix E

Participant informed consent

**Title of Research Project:** The influence of experience and training in the examination pattern of panoramic radiographs: An eye-tracking study

**Investigator(s)**

1. *Principal investigator*
   
   Dr. Daniel Turgeon DMD  
   M.Sc. candidate  
   Division of Oral and Maxillofacial Radiology  
   
   Faculty of Dentistry, University of Toronto  
   124 Edward Street, Division of Oral and Maxillofacial Radiology  
   Toronto, ON M5G 1G6  
   647-876-1541

2. *Faculty Supervisor*
   
   Dr. Ernest Lam DMD, PhD, FRCD(C)  
   Associate Professor  
   Division of Oral and Maxillofacial Radiology  
   
   Faculty of Dentistry, University of Toronto  
   124 Edward Street, Division of Oral and Maxillofacial Radiology  
   Toronto, ON M5G 1G6  
   416-979-4932 x4385

**Purpose of the Research**

The purpose of this study is to evaluate and compare the search patterns used by dental students and oral and maxillofacial radiologists when interpreting panoramic radiographs.
Description of the Research

Subject population

Third year dental students at the Faculty of Dentistry, University of Toronto, and certified (Members or Fellows of the Royal College of Dentists of Canada and Diplomates of the American Board of Oral and Maxillofacial Radiology) oral and maxillofacial radiologists of Canada and the USA.

Inclusion criteria

Students: Participants will be D.D.S. students enrolled in the Faculty of Dentistry of the University of Toronto in the second half of their third year or the first half of their fourth year. Both regular stream students and International Dentists Advanced Placement Program (IDAPP) students will be included.

Oral and maxillofacial radiologists: Participants will be Members or Fellows of the Royal College of Dentists of Canada or Diplomates of the American Board of Oral and Maxillofacial Radiology.

Exclusion criteria

Students: Students would normally have completed an online-seminar in the interpretation of panoramic radiographs during their 3rd year. If a student has not completed the panoramic seminar, he or she will be excluded from this study.

Oral and maxillofacial radiologists: No exclusion criteria.

Time commitment

Approximately 30 to 45 minutes.

Background information

Multiple studies in the medical radiology field have been performed to analyze patterns of search used by radiologists and medical radiology residents when interpreting radiographs. These
studies have led to a better understanding of the way in which radiologist specialists identify abnormalities on radiographs. These findings then have the potential to be used to create different methods of teaching radiologic feature identification and interpretation, using the experts' methods of analysis. In dentistry, little work has been done in this area.

Only one paper has been published regarding eye-tracking for dental images, “Analyzing the eye movement of dentists during their reading of CT images”, by Suwa et al. In this paper, they looked at how eight dental specialists looked at computed tomography (CT) images of the head and neck. Unfortunately, no study has ever looked at a more generalized type of radiograph: the panoramic radiograph.

평가 of the panoramic radiographs

Twenty (20) panoramic radiographs will be shown to each participant. 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, a series of questions relating to this radiograph will have to be answered by the participant. The principal investigator or the supervisor will always be present during these experiments if any problems or questions arise.

Data analysis

The data acquired from each participant will be divided and combined together in their respective participant categories, i.e.: students or radiologists. Multiple parameters, registered by the eye-tracking device will be reviewed and compared to be able to understand the search pattern of both groups.

Potential Harm, Injuries, Discomforts or Inconvenience

No potential harm, injuries, discomforts of inconvenience is expected.

Potential Benefits

Following the completion of the experiment, each participant will be able to request the results of his test.
Confidentiality

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.

Participation

Participation in this research project is voluntary. If you choose to participate in this study, you may withdraw at any time. Should you choose to withdraw, your data will be removed from the database.

Contact

If you have any questions about this study, please contact:

Dr. Daniel Turgeon DMD
M.Sc. candidate
Division of Oral and Maxillofacial Radiology
Faculty of Dentistry, University of Toronto
124 Edward Street, Division of Oral and Maxillofacial Radiology
Toronto, ON M5G 1G6
647-876-1541
daniel.turgeon@dentistry.utoronto.ca

If you have any complaints or concerns about how you have been treated as a research participant, please contact:

Office of Research Ethics
ethics.review@utoronto.ca or 416-946-3273
CONSENT FORM

By signing this form, I agree that:

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>The study has been explained to me.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All my questions were answered.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Possible harm and discomforts and possible benefits (if any) of this study have been explained to me.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I understand that I have the right not to participate and the right to stop at any time.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I have a choice of not answering any specific questions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am free now, and in the future, to ask any questions about the study</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I have been told that my personal information will be kept confidential</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I understand that no information that would identify me, will be released or printed without asking me first.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I understand that I will receive a signed copy of this consent form.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

I hereby consent to participate.

______________________________ ________________
Signature                                              Date

Name of Participant: ____________________________________

Email address: _________________________________________

Name of person who obtained consent: _____________________

____________________________ __________________
Signature                                          Date