The Effectiveness of Systematic Search Strategy Training for the Analysis of Panoramic Images

by

Hanadi M. Khalifa, BDS

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

© Copyright by Hanadi M. Khalifa, BDS 2013
The Effectiveness of Systematic Search Strategy Training for the Analysis of Panoramic Images

Hanadi M. Khalifa, BDS

Master of Science

Graduate Department of Dentistry
University of Toronto

2013

Abstract

Objectives: To examine the effectiveness of systematic search strategy training for the analysis of panoramic images. Methods: 41 students and 33 dentists were recruited from King Abdulaziz University, Saudi Arabia. At baseline, participants analyzed 15 panoramic images, diagnosed whether the images were normal or abnormal, and marked the locations of the abnormalities. One week later, participants in the trained group were shown a training module illustrating a systematic strategy to analyze panoramic images. At post-test, they analyzed 15 new images matched for difficulty with pre-test images. Results: Mixed design ANOVAs revealed that for abnormality detection in the complex part of the panoramic image, trained students improved from baseline but dentists did not. The number of findings over-reported by students and dentists increased after training. Discussion: This study suggests some benefits in the systematic search strategy training for novices. Inclusion of training in normal panoramic features may reduce over-reported findings.
Acknowledgments

I would like to sincerely thank my supervisor, Dr. Michael Pharoah, for his mentorship, encouragement, motivation and support over the past four years.

I would also like to express my deepest appreciation to my co-supervisor, Dr. Vicki LeBlanc, for her expertise in educational research, guidance, and assistance with the statistical analysis that added considerably to my first graduate experience.

To my committee members: Dr. Ernest Lam and Dr. Susanne Perschbacher, I am grateful for your thoughtful contribution and valuable input.

I am eternally grateful to my parents, Fatima and Mohammed, for always being there for me and supporting me throughout my life.

Very special thanks go to my husband, Hussam, for his patience and endless support and to my kids, Asiah and Anas. Thank you for bringing so much happiness in my life.

I would like to extend my sincere appreciation to those students and dentists who volunteered their time to participate in this project. A special thank you goes to the students’ class leader, Ali Arab, for his devoted help and enthusiasm.

I dedicate this work to my sponsor, King Abdulaziz University, Faculty of Dentistry, Jeddah, Kingdom of Saudi Arabia for giving me the opportunity of my life time.
# Table of Contents

Acknowledgments ........................................................................................................... iii
Table of Contents ........................................................................................................... iv
List of Tables .................................................................................................................. vi
List of Figures ................................................................................................................ vii
List of Appendices ......................................................................................................... ix
Chapter 1 .......................................................................................................................... 1

1. Introduction .................................................................................................................. 1
   1.1 Introduction to literature review ............................................................................ 1
      1.1.1 The Process of Radiologic Interpretation ..................................................... 3
      1.1.2 Causes of Errors in Radiologic Interpretation .............................................. 5
   1.2 Summary .................................................................................................................. 16
   1.3 Aim ......................................................................................................................... 16
   1.4 Objectives .............................................................................................................. 17
   1.5 Hypotheses ........................................................................................................... 17

Chapter 2 .......................................................................................................................... 18

2 Materials and Methods................................................................................................. 18
   2.1 Research Design .................................................................................................... 18
   2.2 Sample .................................................................................................................. 18
   2.3 Recruitments ........................................................................................................ 18
   2.4 Testing Materials ................................................................................................. 19
   2.5 Training Material ............................................................................................... 23
   2.6 Instrument .......................................................................................................... 24
   2.7 Research Setting .................................................................................................. 25
   2.8 Research Procedure ............................................................................................ 25
   2.9 Ethical Considerations ......................................................................................... 26
List of Tables

Table 1: The differences between the two sets of panoramic images ............................................. 20
List of Figures

Figure 1: Panoramic image illustrates the subdivision of the image into a complex part above the red line and a simpler part below the red line. ................................................................. 15

Figure 2: Study Sample ............................................................................................................ 19

Figure 3: Panoramic image contains a relatively radiopaque entity in the maxillary left first molar area (asterisk) in the complex portion of the image. ................................................................. 21

Figure 4: Panoramic image contains an obvious radiopaque abnormality attached to the mandibular right third molar (straight arrow) and a subtle widening of the left inferior alveolar nerve canal (asterisks) in the mandibular portion of the image. ......................................................... 21

Figure 5: Mean scores (percentage of correctly diagnosed normal images) of students and dentists at pre-test and post-test. ................................................................................................. 30

Figure 6: Mean scores (percentage of correctly diagnosed abnormal images) of students and dentists at pre-test and post-test. ................................................................................................. 32

Figure 7: Mean scores (percentage of incorrectly diagnosed abnormal images) of students and dentists at pre-test and post-test. ................................................................................................. 34

Figure 8: Mean scores (percentage of correctly detected findings in the complex part of the image) of students and dentists at pre-test and post-test. ................................................................. 36

Figure 9: Mean scores (number of over-reported findings in the complex part of the image) of students and dentists at pre-test and post-test. ................................................................. 38

Figure 10: Mean Scores (percentage of correctly detected abnormalities in the mandibular image) of students and dentists at pre-test and post-tests................................................................. 40

Figure 11: Mean scores (number of over-reported findings in the mandibular part of the image) of students and dentists at pre-test and post-test. ................................................................. 41

Figure 12: Mean Average Confidence (correctly diagnosed normal panoramic images) of students and dentists at pre-test and post-test. ................................................................. 43
Figure 13: Mean Average Confidence (correctly detected abnormal findings) of students and dentists at pre-test and post-test. ................................................................. 44

Figure 14: Mean Average Confidence (over-reported findings) of students and dentists at pre-test and post-test. ................................................................. 45

Figure 15: Mean time (sec) of students and dentists at pre-test and post-test. ......................... 47

Figure 16: Mean scores (percentage of images in which both abnormalities were detected) of students and dentists at pre-test and post-test. ................................................................. 49

Figure 17: The shadow of the soft tissue of the nose (X₁) and the right submandibular fossa (X₂) were mistakenly identified as abnormal findings in this abnormal panoramic image. The true abnormality is in the maxillary left premolar-molar area (dotted arrows). .......................... 53

Figure 18: The incisive foramen (X₁) and the right nasal concha (X₂) were mistakenly identified as abnormal findings in otherwise normal panoramic image. ................................................. 53

Figure 19: The superimposition of the cervical vertebrae over the mandibular midline was mistakenly identified as an abnormal finding in this normal panoramic image. ..................... 54
List of Appendices

Appendix 1: Students Recruitment Message ................................................................. 72
Appendix 2: Experienced Dentists Recruitment Sheet .................................................. 74
Appendix 3: Panoramic Images Evaluation Sheets.......................................................... 75
Appendix 4: The Checklist .............................................................................................. 76
Appendix 5: Written Instructions .................................................................................... 77
Appendix 6: Written Instructions .................................................................................... 78
Appendix 7: Consent Agreement ..................................................................................... 79
Appendix 8: Trained Groups Post-test Questions ........................................................... 81
Appendix 9: Ethics Approval by the Research Ethics Committee of King Abdulaziz
University ..................................................................................................................... 82
Appendix 10: Ethics Approval by Research Ethics Board of University of Toronto ...........81
Chapter 1

1. Introduction

1.1 Introduction to literature review

The radiographic image is a diagnostic tool that plays a significant role in patient care. It provides the clinician with unique information without which diagnostic and treatment decisions often cannot be made. However, radiographic images are not self-explanatory. Therefore, to gain the diagnostic benefit, a qualified practitioner reviews the image and sees, recognizes, and interprets the visual information and then renders a decision about patient management. This process requires a good understanding of how an image has been formed to be able to assign meaning to the features it contains. Additionally, the interpretation task requires a clear understanding of the appearance of normal radiological anatomy and its variations, combined with the ability to discriminate between normal and abnormal appearances.

The process of interpreting a radiograph includes the elements of visual search, detection of abnormal findings, pattern recognition and finally decision-making.\(^1\)\(^2\) One complication of this process is the inherent ambiguity of the radiographic image because it is a two-dimensional image of three-dimensional structures. This results in a significant amount of superimposition of the anatomic structures. Such ambiguity has been reduced by the advent of cross sectional images, such as computerized tomography and magnetic resonance imaging. Nevertheless, radiologic interpretation remains a complex process and has been shown to be subject to errors.

There are reports of a high frequency of errors in radiologic interpretation that may have negative impact on patient care and are a major source of malpractice litigation against radiologists. In a 20 year review of medical radiology malpractice litigation, Berlin\(^3\) disclosed that 70% of all lawsuits filed against radiologists were due to diagnostic errors. The largest proportion of these cases was attributed to the delay in diagnosing malignancies. A review by Goddard et al.\(^4\) found that the range of clinically significant errors across various radiologic investigations, including chest radiology, CT, barium studies and MRI was 2% to 20%. Similarly, Manning et al.\(^5\) found, in a study involving radiologists who examined chest radiographs containing lung nodules, that 27.2% of the lesions were missed. The incidence of
errors, both false positive and false negative, in the interpretation of oral and maxillofacial imaging has not been reported. However, it has been found, in a study by Alsufyani and Lam⁶ to assess the diagnostic differences between three general dentists and three oral and maxillofacial radiologists in the identification of pathognomonic radiographic features of osseous dysplasia and its interpretation, that oral and maxillofacial radiologists identified more features of osseous dysplasia than general dentists and consequently correctly interpreted 79.3% of the osseous dysplasia cases. In contrast general dentists correctly interpreted only 38.7% of the osseous dysplasia cases. Rushton et al.⁷, in a study to identify the radiologic findings in 1,818 screening panoramic radiographs taken for new patients, found that experts (consensus of two oral radiologists) significantly diagnosed more true positive findings than general dentists (n=41). They found that experts identified 366 abnormalities (20.1%), while dentists identified 64 abnormalities (3.5%). Those abnormalities included dental anomalies, jaws abnormalities, both radiolucent and radiopaque entities, TMJ abnormalities and soft tissue calcifications.

These studies demonstrate the difficulty of the radiologic interpretation task, and suggest that a significant number of abnormalities in radiographic examinations may be missed. Errors in radiologic interpretation may have a significant impact on patient care. Incorrect reporting of a non-existing abnormality (false positive errors) may result in healthy individuals having needless treatment. At the same time, missing an abnormality (false negative errors) may jeopardize patient care by delaying the diagnosis and postponing necessary treatment. Errors not only have a detrimental effect on the health and finances of patients but also add a cost to private and public health care systems. In a recent update of radiology liability, Thompson and Patel⁸ state that radiology represents the sixth most frequently sued specialty and the sixth highest in terms of dollars paid in claims. Physician Insurers Association of America claims data from 1985 to 2007 demonstrated 12,970 closed claims against radiologists. Fifty percent of those claims were due to errors in diagnosis, which accounted for 70% of the indemnities paid. Failure to diagnose breast and lung cancer and spinal fractures were the most common conditions associated with errors in diagnosis. Studies indicating the extent to which diagnostic errors involving oral and maxillofacial radiology compromised patients’ care are lacking. Also, no studies document the cost to health care from diagnostic errors involving oral and maxillofacial radiology.
To identify the causes of errors, it is important to understand how an observer analyses the visual information in a radiographic image. In the next section, the process of radiologic interpretation will be discussed.

1.1.1 The Process of Radiologic Interpretation

Understanding the process of radiologic interpretation is of practical importance, as preventing errors depends on our knowledge of the mechanism involved in this process. Radiologic interpretation involves a complex interaction between a practitioner’s visual system and the radiographic image. Many different models have been proposed to explain this process. A useful model for radiologic interpretation was outlined by Blesser and Ozonoff. This model artificially divides the process of data transmission from the radiographic image to the conscious mind, where “diagnosis” occurs, into three phases: psycho-physical, psychological, and nosological.

The psycho-physical phase involves the transformation of a patient’s anatomy into a corresponding two- or three-dimensional radiographic image by the imaging system, and the reception of this visual information through the visual sensory system of the eye. This information is transmitted by the ocular nerve to higher centers in the brain where cognition occurs. In summary, this phase is affected by technical factors of the x-ray source, such as peak kilovoltage and milliamperage, the patients’ physical properties, image processing and display, as well as the human visual system. For example, a practitioner may not see an abnormal finding in a poor quality radiographic image caused by faulty patient positioning or inadequate exposure. As well, an abnormal finding recorded on an optimum image may not be noticed if the image is examined under improper viewing conditions. Alternatively, an abnormal finding on a perfectly exposed image may be overlooked as a result of incomplete visual search.

The second phase of this model, the psychological phase, is an unconscious cognitive process by which the brain attempts to organize the visual data and create a meaningful image or a “visual concept”. The visual concept is formed by data from both the retina and memory and is, in effect, a diagnostic hypothesis about the meaning of the visual data. The diagnostic hypothesis is continuously tested by sampling the image to accept or reject the generated hypothesis. This phase is strongly affected by the expectation of the viewer. For instance, a practitioner may tend to identify some features in the image and/or dismiss others based on what he or she expects to see. Biases created by clinical history, previous experience with similar cases, or laboratory data
may affect the visual concept or even result in the formation of a visual concept before the radiographic image is examined. Also, this phase is affected by knowledge provided by training and experience in radiographic interpretation. Tuddenham\textsuperscript{10} states that “what we perceive is literally as plausible as we can make it on the basis of prior experience, but it is not necessary a faithful representation of the stimulus reaching the retina”. Two practitioners may construct different visual concepts from the same radiographic image. What to one practitioner is a large radiolucent lesion in the posterior mandible, may be a prominent submandibular gland fossa to the other. Thus, previous experience with the visual information is crucial to the formation of the correct diagnosis for the task at hand. This concept was demonstrated in a study by Kundel and Nodine,\textsuperscript{11} who showed a group of laypersons and radiologists a hidden figure picture of a cow and four different radiological images and asked them to give a summary statement about what they saw with recording their eye-movements. They found that radiologists’ verbal responses were in the form of a diagnosis rather than features description. Eye movements revealed that when the image was familiar, visual attention focused on areas related to the generated meaning of the image, whether correct or incorrect. In contrast, with ambiguous images viewers focused their attention on dominant pictorial features, such as borders and texture areas, in order to generate a meaning. Based on their observation, the authors hypothesized that expert visual diagnosis is strongly shaped by the visual concept.

The final phase, the \textit{nosological} phase, (from “nosology”, the taxonomy of disease) is a conscious decision-making phase. During this phase, the learned criteria are used to classify the visual concept and render a diagnosis. This part of the model is influenced by the practitioner’s criteria for “normal” and “abnormal” and on the subsequent deductive clinical reasoning. Therefore, errors may arise due to misclassification of the formed visual concept or faulty reasoning. The reasoning process may become unconscious and faster with training and experience. Thus, a case considered straightforward and diagnosed immediately by a radiologist may require a long and complicated process of deliberate reasoning by an inexperienced practitioner. Smith (as cited in Berlin)\textsuperscript{12} has estimated that faulty reasoning is responsible for 10\% of all errors in radiology. It is important to note that although the occurrence of the three phases appears to be successive and independent of each other, the truth is that they are interrelated.\textsuperscript{9}
In their review of studies of expertise in visual diagnosis, Norman et al. \cite{13} suggest that visual diagnosis in radiology consists of two components: perception is a non-analytical process that involves rapid recognition of patterns in the data and cognition is an analytical process that assesses the output of the perceptual stage and requests further information when needed. Perceptual skills include visual search, visual information processing, and visual discrimination and differentiation between what is relevant and irrelevant.\cite{1} Cognitive skills, on the other hand, are related to diagnostic reasoning and decision making.\cite{1} Errors, therefore, can be classified into perceptual when a practitioner fails to “see” an abnormality, and into cognitive when an abnormality is perceived but incorrectly identified. For example, the same practitioner can miss an abnormality that is evident on a second inspection of the same image (perceptual) or may arrive at different diagnoses of the same radiograph on two separate inspections (cognitive). Renfrew et al.\cite{14}, in an analysis of 182 radiologic errors, found that 60% of the errors were perceptual and cognitive (false negative, false positive, and misclassification).

All of the above factors involved in the process of radiologic interpretation highlight the complexity of the interpretation task and the difficulty in determining the cause of errors. Some of these causes will be discussed in the next section.

1.1.2 Causes of Errors in Radiologic Interpretation

There are many causes for errors in radiologic interpretation. Some of these causes are suggested in the Blesser and Ozonoff\cite{9} model for the radiologic interpretation process. For example, technical factors, such as poor acquisition technique, improper processing or sub-optimum viewing conditions, limit the visibility of an abnormal entity, thus affecting the psychophysical phase of the process. However, even with optimum radiographic images, some abnormalities in the image can be missed. For example, an individual who lacks knowledge about the features of an abnormality that may affect anatomic structures captured in a radiographic image would fail to recognize these abnormalities when interpreting the radiograph; this is a psychological or nosological error. Errors may also be attributed to the visual search, lack of training and experience, the radiographic characteristics and location of the abnormality, and satisfaction of search. These factors will now be explored.
1.1.2.1 Visual search

Visual search is an integral part of the process of radiologic interpretation, and involves two interrelated functions: eye movement and attention.\textsuperscript{1,2,15} The need for a visual search arises from the fact that the retina does not have uniform spatial resolution. The greatest spatial resolution is at the fovea, in the centre of the retina, and resolution diminishes rapidly at the periphery. Thus, in a complex visual image, the image of a small object must be projected closer to the fovea to be perceived. During a visual search, a saccade, a rapid eye movement from one point to another, moves the eye gaze to examine the image. The pause over an area of interest is called a fixation and this average between 200 and 300 milliseconds (msec). These fixations cluster on particular locations in the image to enable the brain to process the visual information projected onto the retina. Thus fixations represent the location of conscious attention. The scan path is the sequence by which the image is sampled, and it gives insights into how individuals prioritise sampling locations. The combination of all these visual activities during a search task produces what is called a search pattern, which is assumed to reflect an individual’s search strategy.

Kundel and Nodine\textsuperscript{16} used a “flash” experiment to investigate how much information radiologists can gain from a single, short glance at a chest radiograph (200msec). Considering that no search was possible in such short duration, they found, surprisingly, that radiologists were capable of detecting 70% of abnormalities present. Detection accuracy improved to 97% when search was permitted for unlimited time. Based on this experiment, Kundel and Nodine developed a global–focal model of visual search and detection. According to this model, search involves global and focal analysis. The global analysis is the initial screening process using the whole retina; it gives the viewer a general impression of the content of the image, including orientation, anatomical layout, symmetry, and overall characteristics. Gross departures or “perturbation” from the normal appearance of anatomy can be flagged at this stage. Input from the global impression subsequently guides the focal analysis, a process called cognitive guidance. During this process, the viewer concentrates the foveal vision on the potential abnormal areas and tests them against the viewer’s mental representation of what is normal. Kundel and Nodine suggest that the total search strategy consists of an ordered sequence of interspersed global and checking fixations, and that visual search has both perceptual and cognitive components.\textsuperscript{1,2,15}
Visual search has been studied extensively through the use of eye-position tracking devices. The goal of these studies is to understand the components involved in the development of expertise and to identify why errors tend to occur. These studies\textsuperscript{10, 17, 18} have shown that incomplete visual search is a potential source of false negative errors. For instance, Tuddenham and Calvert\textsuperscript{19} were the first to record the visual search patterns of radiologists who were presented with a series of radiographs of normal and abnormal chests. The radiologists scanned the images with a movable spotlight while a motion picture camera recorded the locations of the spotlight on the images. The authors found that there was incomplete coverage of the image by the spotlight, and speculated that an inadequate search of the images might be a significant source of errors.

Kundel et al.\textsuperscript{17} followed up on these ideas and used eye-position tracking to analyse the eye movements of three radiologists and one inexperienced observer searching for a simulated nodule in chest radiographs. Based on their research data, the authors classified the false negative errors into three categories: search errors (30\%), pattern recognition errors (25\%), and decision-making errors (45\%). The authors believed that search errors occurred because of incomplete searches of the radiographic image and failure of radiologists to fixate close enough to the missed lesions. Pattern recognition errors occurred when missed nodules were fixated on for a short period of time, and these were never returned to. Subjectively, it appears that the observers fail to recognize the nodule features as abnormal. Decision-making errors were considered when the eyes fixated on the lesions for a long period of time but the lesions were not reported. The same breakdown of errors has been noted in missed fractures in skeletal radiographs\textsuperscript{18} and missed masses in mammography images.\textsuperscript{20} One limitation with this classification system is that the difference between pattern recognition and decision errors is based on a specified fixation threshold time. This threshold time may not apply for all abnormalities seen on different types of radiographic images.

To reduce errors, students, in both dentistry and medicine, have been advised to search the radiographic image in a systematic manner to achieve a complete examination of all the anatomical regions in that image. Systematic search refers to the visual examination of the radiographic image in a specific order by directing the observer's attention to a sequence of anatomical regions within the radiologic image, sometimes referred to as guided or focused search. For instance, Worth\textsuperscript{21} emphasizes that "the examiner must give his undivided attention to the study of every part of the radiograph. It is wise for him to adopt some orderly system which
will ensure that every portion of every film comes under study". Sutton (as cited in Carmody et al.\textsuperscript{20}) states that “a film of the chest is too large to be taken at a glance and radiologists must, therefore, develop a technique for examining the film systematically”.

A systematic search strategy should be reflected in eye-movement behaviour but in truth, it is not. Evidence from published eye-position tracking experiments using chest radiographs demonstrates that most experienced radiologists do not adhere to any systematic strategy\textsuperscript{19, 22}. In the previously mentioned Tuddenham and Calvert study\textsuperscript{19}, it was noted that there was a large intra- and inter-observer variability in radiologists’ search patterns, although all of them were trained in the same department. One radiologist employed a systematic pattern, another used an initial circumferential pattern followed by a more complex scan, and the other two used an unclassified search pattern. With respect to errors, the one radiologist who used a systematic search missed significant findings that were noted by the other observers. The authors suggest that adopting a consistent systematic search pattern has a detrimental effect on radiologic interpretation performance. A weakness in this study was the low number of observers. Carmody et al.\textsuperscript{22} found there to be a discrepancy between instruction and actual radiograph reading practices. They surveyed radiologists and residents to investigate how they were taught to examine chest radiographs. Eye-positions were then recorded in a different group of radiologists and residents as they examined chest radiographs. This survey showed that radiologists were taught to examine chest radiographs systematically with comparing of features on both sides of the image. However, the eye movement experiment revealed that 83\% of the participants did not follow the systematic search and that comparisons of bilateral features comprised less than 4\% of the visual search. These workers found that radiologists focused their attention on similar anatomic structures; however, the sequence of fixation on particular area on the images was not consistent, resulting in different scanning patterns. One major criticism of this study is that the survey and the eye-tracking experiment were conducted on different groups of radiologists which could account for the reported differences between training and practice. Contrary to the observation by Carmody et al.,\textsuperscript{22} Hu et al.\textsuperscript{18} observed a more systematic scanning pattern for experienced radiologists than inexperienced observers when examining hand-wrist radiographs for bone fractures. Given these findings, one might hypothesize that the type of search used by the radiologist can be influenced by the complexity and anatomical layout in various types of radiologic images.
Based on their findings, Carmody et al.\textsuperscript{22} advocate the use of a systematic strategy as a “pedagogical aid”, useful for training novices, but not experienced radiologists, since this leads to better and more complete appreciation of the appearance of variations of normal anatomy. They argue that experts drop the systematic strategy as soon as they know the range of normal variations. With experience, practitioners have a better ability to recognize the range of variation of normal features and discriminate abnormal features from normal. According to Carmody et al.,\textsuperscript{22} an expert’s search follows the global-focal model in which the expert’s search strategy is established by background knowledge about the image and its anatomic structure. The search is subsequently modified by input from the global impression. Lesion-free images are analysed in a systematic manner while images that contained abnormal findings received focal attention directed to those areas flagged at the global impression. These latter areas are then compared with background features found in other parts of that image. They state that “the comparisons are a result, not a cause of detecting variations beyond the range of normal”. Taylor\textsuperscript{23} came to the same conclusion, in his review of research into the development of radiologic expertise. He states that “the impact of the training is not that readers learn the right way to move the eyes across the image but that training provides a cognitive basis for recognizing ‘perturbation’ deviations from the radiologist’s mental model of normal appearance”.

The utility of a systematic search strategy in the analysis of chest radiographs has been challenged in the medical literature. One attempt to force medical students to adhere to a specific systematic search strategy did not improve their overall diagnostic accuracy. Gale and Worthington\textsuperscript{24} conducted an experiment to test whether students trained to use a systematic analysis of chest radiographs would perform better. The performance measure was the accuracy in detecting nodules in chest images. Students were divided into experimental and control groups. The experimental group was shown a training videotape to encourage them to adopt a specific systematic search strategy, and the control group received no such training. The results showed no statistically significant differences in the performance of the control and experimental groups. The authors attributed the lack of significant improvement in the performance of the experimental students to the dual-task demands of examining the radiograph and adhering to the systematic strategy. They speculated that these two tasks may compete for attention allocation. This explanation was supported by a significant increase in the response time of the experimental group for both true positive and true negative findings. The authors concluded that the adoption
of a systematic search, although simple and requiring less memorization, does not improve observers’ performances. As well, the authors believed that the use of a systematic search as prescribed in the radiological texts and manuals is not evidence-based.

Swensson et al. compared the performances of radiologists under two different search conditions: normal examination of the entire film (free search) and examination of specified regions on chest radiograph (systematic search). Some of the regions were broadly specified, such as upper lung fields, abdomen, soft tissue and bone. Other regions were narrowly specified, such as cardiac border and retro-cardiac region. Their hypothesis was that a systematic search would eliminate false negative errors because it directs radiologists’ attention to all critical anatomical features in the radiograph. They found a large increase in the number of both true positive and false positive reports with the systematic search. Based on their data, they suggest that, rather than improving the radiologists’ ability to discriminate normal from abnormal, the systematic search lowers participants’ threshold to identify a feature as abnormal. In other words, the systematic search made participants more willing to report questionable findings as abnormal when directed to focus on them.

Another attempt to impose a checklist to examine chest radiographs in a particular order also resulted in poor performance of participating radiologists. Berbaum and Franken found that radiologists performed poorly when they were asked to follow a checklist when analysing chest radiographs. The findings were different when radiologists were asked to search chest radiographs verbally by indicating what structures they were looking at while simultaneously pointing to the image and what features they looked for as they searched the image. Berbaum and Franken found that most observers’ verbal reports demonstrated an often personally unique, but deliberate search over many chest radiographs. In terms of performance, the accuracy of detection of native abnormalities when a simulated nodule was added was higher with a verbal search than with the use of a checklist. The explanation proposed by the authors is that the checklist interrupted observers’ internal strategy of searching radiographs. Although the manipulation in these studies differed, they demonstrate that performance in the analysis of chest images suffers when a systematic search strategy is followed. Thus, the results of these studies challenge the use of a systematic search strategy with novices and experts.
It is necessary to note that most of the previous referenced studies used chest radiographs as the test material. One question that needs to be asked, however, is whether these findings are applicable to other types of images and specifically panoramic images, which are commonly used in dentistry. In a retrospective analysis of referral pattern for oral radiologic consultation in Ontario\textsuperscript{28}, panoramic radiographs, alone or in combination with other film types, were submitted for consultation in 79.5\% of cases. This number may reflect the greater use of this image in dental practice. Also, this number may indicate that a panoramic image is a frequent source of interpretation uncertainty due to its complexity. Rushton et al.\textsuperscript{29} conducted a survey to determine the factors influencing the prescribing of panoramic radiographs in general dental practice and found that 79\% of dentists were inclined to acquire a panoramic radiograph at the initial visit.

Given the differences between chest and panoramic images, it is likely that the research findings using chest radiographs are not applicable to panoramic images. Despite the smaller size of the panoramic image, it encompasses a high concentration of complex osseous structures of the entire maxillofacial region including the dentition, mandible, maxilla, parts of the maxillary sinuses and nasal structures, parts of the sphenoid and temporal bones, as well as parts of the cervical vertebrae. Therefore, there are more anatomic features per unit area that need to be inspected compared to chest radiographs. In addition, a panoramic radiograph is considered a specialized application of tomography and represents an image of a curved image slice of anatomy obtained by selective blurring. This acquisition technique results in a complex image due to the superimposition of the hard and soft tissue structures of the region, variable distortion, ghost shadows, and low spatial resolution, which make visualization of fine details and subtle diseases difficult. In contrast, the chest radiograph is a conventional projection radiograph obtained directly without selective blurring. Due to the complexity of the panoramic image, it is possible that a systematic search strategy would be more successful in the detection of abnormalities and likely the reason why this is a method of choice in training novice students. In addition, considering the fact that panoramic radiographs are commonly prescribed for new patients for the purpose of dental diagnosis and treatment planning, a dentist may fixate on the dentition at the risk of neglecting the rest of the image. Also, a patient’s chief complaint, or signs and symptoms may bias the dentist and draw his or her attention to a particular area on the image, whereas other subtle abnormal findings can be missed. Given that not all diseases have noticeable signs and symptoms, one should consider the importance of a careful and consistent
systematic search strategy to ensure that all anatomic structures in the image are indeed read. It remains unclear, however, whether training either students or experienced dentists in a systematic search strategy for the analysis of panoramic images would improve their radiologic interpretation performance.

1.1.2.2 Lack of Training and Expertise

Training and experience play a role in the development of the perceptual and cognitive skills in radiologic interpretation.\textsuperscript{1, 2} Kundel and Le Follette\textsuperscript{30} suggest that accuracy and expertise is based to a large extent on the perceptual process of the diagnostic task. They used eye-position tracking to analyze the search pattern of participants with different levels of experience (from lay-people to radiologists) while viewing normal and abnormal chest radiographs. The study demonstrated that novices and experienced observers differ in their visual search. In their study, the visual search ranged from a localized central search by first-year medical students to a circumferential search by experienced radiologists. Also, the authors found that experience fine-tunes the visual recognition of abnormalities. Inexperienced observers fixate on a lesion but do not recognize the abnormality. In contrast, experienced observers fixate on and report abnormalities correctly. Similarly, Krupinski\textsuperscript{20}, using mammograms as the test material, found that radiologists were both faster and more accurate than residents in interpreting mammograms. In Krupinski’s study, radiologists tended to detect lesions earlier in the search, and when multiple lesions were present in an image, radiologists detected all lesions rapidly in a single glance. In contrast, the less experienced residents tended to deliberately search the image between successive detections and were less likely to recognize normal from abnormal. Kundel et al.\textsuperscript{31} assume that experts develop a “holistic recognition” skill, therefore, their search could be modified from a search to find strategy to a more efficient one, similar to one used for facial recognition. Myles-Worsley et al.\textsuperscript{32} suggest that radiological expertise is based on two main bodies of knowledge: knowledge of features of clinically normal examples and knowledge of the set of abnormal features that signal pathoses.

Given these observations, it is possible that experienced dentists who have analyzed more panoramic images are familiar with the normal anatomy of the panoramic radiograph and have previously encountered more pathology. In doing so, these experienced dentists may have developed an efficient and unique search strategy and have a superior ability to recognize abnormal features on a panoramic image. In contrast, students are more likely to encounter
difficulties in this task. Thus, it seems likely that training novices in a systematic search strategy would be of benefit, but training more experienced dentists in the same way might not improve their radiologic interpretation performance.

1.1.2.3 Satisfaction of Search

In a retrospective analysis of findings in 14 patients with overlooked lung cancers at CT imaging\textsuperscript{33}, it was noted that 43\% of these image sets had major distracting findings elsewhere in the thorax, such as aortic dissection, aneurysm, and esophageal cancer. These findings might have diverted radiologists’ attention from the missed cancer. Tuddenham\textsuperscript{10} hypothesized that false negative errors may arise from premature termination of the search due to detection of a significant finding or a finding that answers a clinical question, a phenomenon called "satisfaction of search" or “satisfaction of meaning”. Berbaum et al.\textsuperscript{34} were the first to experimentally verify the satisfaction of search phenomenon. They found that adding a simulated nodule (distractor) to chest radiographs that already contained subtle native abnormalities diminished the accuracy of the detection of the native abnormality. Ashman et al.\textsuperscript{35} compared the detection rates in skeletal radiographs containing a single abnormality with those with multiple abnormalities. They found that the detection rate for the second and third abnormalities in the multiple abnormalities cases was reduced to about half. The explanation for this phenomenon is controversial. Tuddenham\textsuperscript{10} speculated that search is terminated after detecting an abnormality. This explanation has not been supported by subsequent investigation.\textsuperscript{34} Eye-position analyses showed that observers continued to search images after reporting an initial abnormality and most missed lesions were fixated on but then ignored. Barbaum et al.\textsuperscript{34} hypothesized that satisfaction of search is related to “perceptual set effects”. Barbaum states that “concluding that certain image features are indicative of one diagnostic category might make it difficult to detect features of another diagnostic category”.

Although the satisfaction of search phenomenon is not simply due to a faulty search, it would be intriguing to examine if training students and experienced dentists in a systematic search strategy decreases the number of false negative errors in panoramic images with multiple abnormal findings.
1.1.2.4 Characteristics of the abnormality

Errors may also occur due to characteristics of the abnormality itself, such as a margin’s definition, and abnormality’s density or location. This was demonstrated in a study by Austin et al.\textsuperscript{36}, who retrospectively analyzed the findings in 27 patients with overlooked bronchogenic carcinoma on chest radiographs. The authors found that blurred definition of the margin and slight level of opacity were characteristics of the missed cancers. Also, 81\% of the lesions were located in the upper zones of the lung. The authors attributed the high frequency of missed lesions in the upper lung, in their cases, to the presence of overlapping osseous structures.

The imaging characteristics of lesions that affect the maxillofacial structure are variable and it is likely that their perception may be influenced by their location, size, shape, density, and behaviour. A study by Stheeman et al.\textsuperscript{37} found that dentists were more accurate in detecting the presence of radiolucent abnormalities than radiopaque ones. In contrast, Perschbacher et al.\textsuperscript{28} in a retrospective analysis of referral patterns to oral radiologic consultation in Ontario, observed that radiopaque entities were submitted for consultation more frequently (37\%) than radiolucent or mixed radiolucent-radiopaque entities. Perschbacher et al. explained this observation by speculating that radiopaque entities, which stand out relative to the radiolucent background of trabecular bone, could be more easily perceived than radiolucent ones. Perschbacher et al., however, did not examine whether maxillary or mandibular findings were submitted more frequently. This data may give insight on the effect of location on abnormality detection.

In the panoramic image, the upper portion containing the orbit, skull base as well as nasal, maxillary and zygomatic bones is complex because of the greater amount of superimposition of hard and soft tissue structures in the area. In contrast, the lower part, containing the mandible, is simpler (Figure 1). Considering the difference in the complexity of these two regions it would be of interest to know if training practitioners to utilize a systematic search strategy has a differential effect on the correct detection of abnormalities in the complex versus simple section of the image.
Figure 1: Panoramic image illustrates the subdivision of the image into a complex part above the red line and a simpler part below the red line.
1.2 Summary
Several published studies in the medical literature have challenged the effectiveness of using a systematic search strategy for the analysis of chest radiographs. However, these findings may not be applicable to the detection task in panoramic images, as both the complexity and display of the anatomic structures are substantially different. Investigation of the effectiveness of search strategies training for the analysis of oral and maxillofacial radiographic images is lacking.

1.3 Aim
This study will test the effectiveness of systematic search strategy training on the radiologic interpretation performance of novices and experienced dentists. For the purpose of this study, radiologic interpretation performance means the diagnosis of the panoramic image as normal or abnormal, and the ability to detect abnormalities. Also, of interest is whether the systematic search strategy training helps both novices and experienced dentists to detect multiple abnormalities in the image. If the systematic search strategy training proves to be effective, inclusion and implementation within undergraduate and graduate curricula could be substantiated. If it proves to be effective for experienced dentists, then inclusion in continuing education courses could be considered. The underlying aim is to reduce errors in the interpretation of panoramic images in order to achieve positive patient outcomes.
1.4 Objectives

The objectives of this study are to:

1. Examine whether training in a systematic search strategy for the analysis of panoramic images improves the radiologic interpretation performance of novices (dental students) and more experienced observers (dentists);

2. Determine whether systematic search training improves the detection of multiple abnormalities in panoramic images that contain multiple abnormal findings.

1.5 Hypotheses

1. A. Training dental students in a systematic search will improve their radiologic interpretation performance with panoramic images;

1. B. Training experienced dentists in a systematic search strategy will not improve their radiologic interpretation performance with panoramic images.

2. Training novices and experienced dentists in a systematic search strategy will improve the detection of multiple abnormalities in panoramic images that contain multiple abnormal findings.
Chapter 2

2 Materials and Methods

2.1 Research Design

This study was a quantitative, randomized, two groups pre-test and post-test design. It measures the effect of training in a systematic search strategy on radiologic interpretation performance of novices and experienced dentists with regards to the correct diagnosis of panoramic images and detecting abnormalities in these images.

2.2 Sample

A convenience sample of 80 volunteers participated in this study. The sample consisted of two main groups: novice dental students and experienced dentists. The novice group consisted of fifth-year dental students from King Abdulaziz University (KAU), Jeddah, Saudi Arabia. Fifth-year dental students were selected because they have completed the oral radiology courses, which are given in the third and fourth years. The more experienced group consisted of dental specialists, who were faculty members at KAU, and general dentists, who were joining the Saudi Board speciality programs at KAU. To be eligible for the study, participants in the experienced group had to have been in practice for a minimum of five years and the analysis of panoramic radiographs was a routine part of their practice. Forty-four students and 36 dentists agreed to participate in the study.

The novice group consisted of 21 (47.7%) males and 23 (52.3%) females, and the experienced dentist group consisted of 14 (42.9%) males and 22 (61.1%) females. Six participants, three students and three dentists, dropped out of the study after the pre-test, leaving a total of 74 participants. (Figure 2)
2.3 Recruitments
Recruitment was made by posting an electronic recruitment message on a students’ electronic message board (Appendix 1). Students were asked to email the principal investigator if they were interested in the study. To recruit the experienced dentists, the principal investigator spoke with potential participants on a one-on-one basis. Each participant was given a recruitment sheet that briefly explained the study and his/her voluntary participation was requested (Appendix 2).

2.4 Testing Materials
The testing materials were two sets of digitized panoramic images selected from the patient database and teaching archives of the Discipline of Oral and Maxillofacial Radiology at the University of Toronto. All patient identifiers were removed from the images. Each image set had a total of 15 panoramic images, with seven normal images and eight contained one or two abnormalities. Some of these abnormalities were within the complex part of the image and others were within the mandibular image (Figure 1), some were relatively radiopaque and others were relatively radiolucent, and some were “obvious” and others were “subtle”. In an effort to decrease pre-test to post-test learning effects, different abnormal panoramic images, matched in difficulty to the pre-test images were used at post-test. The same normal panoramic images were used for both tests; however, the post-test images were flipped in the horizontal plane. The differences between the pre-test and post-test image sets are shown in Table 1. Figures 3 and 4 below are examples of the testing materials.
Table 1: The differences between the two sets of panoramic images

<table>
<thead>
<tr>
<th>Features</th>
<th>Pre-test set</th>
<th>Post-test set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of abnormalities within the complex image</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Number of abnormalities within the mandibular image</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>Total number of abnormalities</td>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td>Number of images with two abnormalities</td>
<td>2</td>
<td>5</td>
</tr>
</tbody>
</table>
Figure 3: Panoramic image contains a relatively radiopaque entity in the maxillary left first molar area (asterisk) in the complex portion of the image.

Figure 4: Panoramic image contains an obvious radiopaque abnormality attached to the mandibular right third molar (straight arrow) and a subtle widening of the left inferior alveolar nerve canal (asterisks) in the mandibular portion of the image.
Selection of the Testing Materials

Two groups of reviewers, three general dentists and three certified oral and maxillofacial radiologists, independently evaluated 37 panoramic images prepared in a PowerPoint® (Microsoft Corp., Redmond, WA) presentation. Each reviewer was asked to determine whether the panoramic image was normal or abnormal (containing at least one abnormal finding), the number of abnormalities present and ranked the difficulty level of each abnormal finding from 1 to 10, with 1 as “very subtle” and 10 as “very obvious”. An obvious finding was defined as an abnormality that can be easily and quickly identified—without any searching. Reviewers’ responses were recorded on sheets (Appendix 3). The principal investigator then calculated the average ranking of each abnormal finding.

The selection criteria of the testing materials were as follows:

- Normal panoramic images must have 100% agreement amongst reviewers that they are normal.
- The abnormal panoramic images should contain abnormal findings that were neither very subtle nor very obvious (i.e., having a difficulty level of between 4 and 7).
- The abnormal panoramic images with two abnormalities should contain one obvious finding (i.e., having a difficulty level of between 8 and 10) and another more subtle (i.e., having a difficulty level of between 4 and 7). An obvious finding was included to test for the influence of the obvious finding on the detectability of the more subtle one.

Inter-rater reliability was determined by an intraclass correlation coefficient. The average measure was 0.78, suggesting that there was “strong agreement” between reviewers. The final selection of the two sets of testing materials was done by the principal investigator. The images in the two sets were chosen to be matched in difficulty. The mean difficulty level (± SD) was 7.09 ± 1.66 and 7.42 ± 1.7 for the pre-test and post-test image sets, respectively. The two image sets were also matched in content in term of abnormalities location and features, such as widened inferior alveolar nerve canal, radiolucent lesions within the mandible, disruption in the continuity of the floor of the maxillary sinuses, lesions within the maxillary sinus, and condylar lesions.
2.5 Training Material

The training material was the Dental Procedure Education System (DPES) module authored by Dr. Susanne Perschbacher and Dr. Mindy Cash, who are faculty members in the department of Oral and Maxillofacial Radiology at the University of Toronto. The material was produced through collaboration with the University of Toronto, Faculty of Dentistry’s Media Services. This module is an audio-video recording combined with panoramic images and visual illustrations. The video is six minutes in duration and illustrates a stepwise methodological approach to the analysis of the panoramic image. The training video cued participants to assess the osseous structures and surrounding soft tissues, examine the alveolar processes, and finally evaluate the teeth. The systematic search strategy outlined by Perschbacher\textsuperscript{38} includes:

1. Assessment of the periphery and corners of the image. Structures that may be seen in this area include: the orbits, articular processes of the temporal bones, cervical spine, styloid processes, pharynx and hyoid bone.

2. Examination of the outer cortices of the mandible by tracing the periphery of the bone starting at one spot and completing a circuit which includes: anterior and posterior rami, coronoid processes, condyles and condylar necks and inferior border and to look for continuity and evenness of the cortices.

3. Examination of the cortices of the maxilla. This includes the posterior, medial walls and floor of each maxillary sinus. Also to look at the zygomatic process of the maxilla and the pterygomaxillary fissure

4. Examination of the zygomatic bones and arches.

5. Assessment of the internal density of the maxillary sinuses.

6. Assessment of the structures of the nasal cavity and the palates.

7. Examination of the bone pattern of the maxilla and mandible and assessment of the density and pattern of the trabeculae for abnormalities.

8. Examination of the size, position, cortication and symmetry of the inferior alveolar nerve canals, mandibular foramina and mental foramina.
The training video was enforced with a checklist that was given to each participant in the trained groups to be used while viewing the images. The purpose of including the checklist was to encourage participants to follow the sequence provided in the video (Appendix 4).

### 2.6 Instrument

Three versions of a computer program were created using Adobe® Flash® Professional CS5.5 (Adobe Systems, San Jose, CA) that contained a pre-test, a post-test for the trained groups and a post-test version for the control groups. The three versions differed in their test instructions and panoramic images. In addition, the post-test version for the trained groups contained the training video and the post-test questions. The program path into the different versions was determined by the entered identification data of each participant. The program started out by asking participants to enter the identification number, session number (1 or 2), group (A or B) and their educational-level (student or dentist). The program then presented the test instructions, which were the same for all participants at pre-test (Appendix 5) but differed between the control and trained groups at post-test (Appendix 6). The post-test version for the trained groups automatically played the training video after presenting the test instructions. Trained participants were able to stop the video for note-taking purposes, and the video had to be completed for the program to advance to the test images.

The test portion of the program started out by asking the participant to determine if the panoramic image was normal or abnormal. If the participant determined that the panoramic image was normal, the participant was presented with a “confidence scale” and asked to place a mark indicating the degree of confidence in their decision. This scale was a visual analogue scale defined by two end-points, “suspicious” and “definite”, where suspicious equalled zero and definite equalled ten. The computer program assigned a discrete value from 0 to 100 to the mark placed on the linear scale by the participants. If the participant determined that the panoramic image was abnormal, he or she was asked to mark the abnormality directly on the image with an “X” symbol, and then was presented with the same confidence scale. The participant had the opportunity to mark more than one abnormality and was presented with the confidence scale after each detected abnormality. The computer program also had the following features: it presented all images in a random order to all participants, recorded the time taken to identify
whether the panoramic image was normal or abnormal, and it did not permit participants to go back to the previous images.

After test completion, the program generated a file for each participant that contained all of the captured images plus a WordPad file for other test responses. The participant’s ID number was the unique filename and the documents were automatically saved to a folder called “Radiology tests” on the computer desktop.

To evaluate the functionality of this computer program, it was tested with faculty members and residents in the Discipline of Oral and Maxillofacial Radiology at the University of Toronto. The program was adjusted based upon constructive input from the reviewers. A pilot test was completed with the participation of four oral radiology residents in order to ensure that it could be used easily and was free of technical glitches.

2.7 Research Setting

The setting for this research study was in the Faculty of Dentistry at King Abdulaziz University. Due to participants’ various schedules, their clinical and courses requirements, participants took the tests individually or in groups, according to their availability. The images were viewed using a 17-inch widescreen LCD Monitor (HP Compaq LE1711, Hewlett-Packard Corporation, Palo Alto, CA) with a display resolution of 1024 x 768. The ambient lightening in the viewing room was subdued.

2.8 Research Procedure

Participants enrolled in the study participated in two tests: a pre-test and a post-test. Both the pre-test and post-test scores were used in the data analysis to measure potential changes in test scores after the training. The two tests were separated by a period of one week in order to control for extraneous variables, such as the effect of repeated exposure to the same task, which could affect the study outcome. In an effort to decrease the pre-test- post-test learning effects, testing materials in the pre-test and post-test were not identical, but were matched in difficulty and content. Informed consent was obtained from each participant prior to participation (Appendix 7). Participants were given ID numbers which were used to code their pre-test and post-test data files in order to keep their identity concealed. The students were randomly assigned to either control (A) or trained groups (B), which consisting of 22 students each. The dentists were also
randomly assigned to either control (A) or trained groups (B), consisting of 18 dentists each. Participants’ allocation to either control or trained groups was achieved by using computer-generated random numbers. Each participant was assigned to a computer and given verbal and written instructions about the study (Appendix 5).

At pre-test, all participants viewed the same set of 15 panoramic images with no instruction provided about using a specific search strategy. Participants were instructed to search for abnormal findings that they thought might require consultation, further investigation or treatment. Also, participants were specifically told not to report caries, periodontal bone loss, and impacted and missing teeth. In addition, they were informed that the image set contained ‘normal’ (i.e., abnormality-free) and ‘abnormal’ (i.e., abnormality-containing) images, but not the proportions of these image types. No time limit was set for viewing, however, the time taken to determine whether the image was normal or abnormal was recorded. No clinical information was provided.

Immediately after test completion, the program presented the participants in the trained groups with two multiple choice questions, about the training video and the use of a checklist in the analysis of panoramic images (Appendix 8). Following the completion of both tests, the principal investigator reviewed the correct diagnoses and the location of the abnormal findings with participants. In this debriefing period, the aim of the study was explained, and participants in the control groups were offered the opportunity to view the training video.

2.9 Ethical Considerations

Approval for conducting the study was granted by a Research Ethics Committee of KAU (Appendix 9) and the Research Ethics Board of the University of Toronto (Appendix 10). The participants were informed that their participation was voluntary and that they could withdraw from the study at any time without incurring penalties. The results were kept confidential, and participants’ anonymity was protected. The ID numbers sheets were destroyed afterward, with no record of how the numbers matched the participants. All data were transferred into a password-protected computer and will be stored for five years following the publication of the results and then destroyed. All data files were permanently deleted from the laboratory computers.
2.10 Data Analysis

The study aims to analyze the effectiveness of systematic search strategy training in the analysis of panoramic images upon the radiologic interpretation performance of students and dentists. In this study, the researchers were interested in 11 radiologic interpretation performance measures. The first three were measures of the diagnostic accuracy of panoramic images. The outcome variables computed were (1) the percentage of correctly diagnosed normal panoramic images; (2) the percentage of correctly diagnosed abnormal panoramic images; and (3) the percentage of incorrectly diagnosed abnormal panoramic images. Four performance measures were related to the abnormality detection accuracy. The outcome variables computed were (4) the percentage of correctly detected abnormal findings in the complex part of the image; (5) the percentage of correctly detected abnormal findings in the mandibular image; (6) the total number of over-reported findings in the complex part of the image; and (7) the total number of over-reported findings in the mandibular image. The other performance measures were related to confidence levels. For each participant the average confidence level was calculated as a value from 0-10 for the following: (8) correctly diagnosed normal panoramic images, (9) correctly detected abnormal findings, and (10) over-reported findings. The last performance measure was (11) search time.

For each outcome variable, separate scores were calculated for the pre-test and post-test. The effect of the training was determined by measuring the difference between the pre-test and post-test scores. To test the effectiveness of the systematic search strategy training on the diagnostic accuracy, abnormality detection accuracy and confidence of students and dentist, each dependent variable was subjected to a mixed design analysis of variance (ANOVA), with participants’ educational levels (experienced dentists vs. students) and groups (control vs. trained) as between-subject variables and time (pre-test vs. post-test) as a within subjects variable. Post-hoc analyses were conducted with paired t-tests and/or independent samples t-tests, as appropriate.

To analyze the effect of the training on the participants’ performance with respect to the detection of both abnormal findings in the five panoramic images with multiple abnormalities (two per image) in the post-test set, each participant was given a score which represented the percentage of the images in which both abnormalities were detected. These scores were subjected to a 2 X 2 (groups X educational levels) factorial ANOVA. All statistical analyses were performed using IBM® SPSS® Statistics 20 software. Differences between groups were
determined to be significant when $p<0.05$. The responses from the post-test questions were analyzed using descriptive statistics.
Chapter 3

3 Results

This section summarizes the research findings related to each of the two research objectives and the post-test questions results.

3.1 Objective One

The first objective was to test whether systematic search strategy training improves students’ and dentists’ radiologic interpretation performance which includes, diagnostic accuracy of the panoramic images, abnormality detection accuracy and confidence level.

3.1.1 Diagnostic Accuracy of Panoramic Images

The diagnostic accuracy of the panoramic images was calculated as a score equal to the percentage of correctly diagnosed normal panoramic images (out of 7) and the percentage of correctly diagnosed abnormal panoramic images (out of 8) at pre-test and post-test. The diagnostic accuracy of the normal panoramic images was coded as: 0 when normal called abnormal; and when normal called normal. The diagnostic accuracy of abnormal panoramic images was coded as: 0 when abnormal called normal; 1 when abnormal called abnormal and at least one of the detected finding was correct; and 2 when abnormal called abnormal and none of the detected findings were correct. Since the researcher was interested in correct diagnosis of panoramic images and correct detection of abnormal findings, responses coded 2 were excluded from the data analysis.

3.1.1.1 Correct Diagnosis of Normal Panoramic Images

The ANOVA of the mean percentage of correctly diagnosed normal panoramic images (calling a normal image normal) revealed no main effect of time, $F(1, 70) = 0.599, p = 0.442$. There was a main effect of educational levels, with dentists outperforming students, $F(1, 70) = 9.521, p = 0.003$. There was also a main effect of groups, with control groups outperforming training groups, $F(1, 70) = 18.558, p < 0.05$. The time by groups interaction was not significant, $F(1, 70) = 0.216, p = 0.643$, neither was the time by educational levels interaction, $F(1, 70) = 3.695, p = 0.05$, the groups by educational levels interaction, $F(1, 70) = 0.0, p = 1.0$, or the time by groups
by educational levels interaction, $F(1,70) = 0.029, p = 0.865$. As shown in Figure 5, neither students nor dentists improved after the training in correctly diagnosing normal panoramic images. For both the dentists and the students, scores for the control groups were higher than the scores for the trained groups, both at pre-test and at post-test.

![Figure 5: Mean scores (percentage of correctly diagnosed normal images) of students and dentists at pre-test and post-test. Error bars represent standard error. $p < 0.05$]
3.1.1.2 Correct Diagnosis of Abnormal Panoramic Images

The ANOVA of the mean percentage of correctly diagnosed abnormal panoramic images (calling an abnormal image abnormal and at least one of the detected finding was correct) revealed a main effect of time, $F(1, 70) = 82.017, p < 0.05$, indicating that overall, participants made more correct diagnoses on post-test than on pre-test. The dentists made more correct diagnoses than students, as shown by a significant main effect of educational levels, $F(1, 70) = 31.963, p <0.05$. Overall, there was no main effect of groups, $F(1, 70) = 2.216, p = 0.14$. However, the time by groups interaction was significant, $F(1, 70) = 11.932, p = 0.001$, indicating that groups performed differently on pre-test and post-test. There was no significant time by educational levels interaction, $F(1, 70) = 0.048, p = 0.826$, nor an educational levels by groups interaction, $F(1, 70) = 0.342, p = 0.56$. The three-way time by groups by educational levels interaction was significant, $F(1 ,70) = 6.86, p = 0.011$, indicating that performance of the control and trained groups differed between time, and this difference was not similar for both educational levels. As such, subsequent 2-way mixed design ANOVAs, conducted separately on dentists and students’ scores, revealed that the time by groups interaction was significant for students, $F(1, 39) = 21.108, p <0.05$, but not for dentists, $F(1, 31) = 0.307, p = 0.584$.

To analyze the effect of time, post-hoc analyses with paired samples $t$-tests were performed on the students’ scores on pre-test and post-test. There was a significant difference between the pre-test and post-test scores for the trained students, $t (20) = -7.408, p <0.05$, but not for the control students, $t(19) = -1.63, p = 0.119$, indicating that trained students significantly improved from pre-test to post-test. To analyze the effect of groups, independent samples $t$-tests were performed on control and trained students’ score at the pre-test and post-test, separately. The groups were not different from each other at pre-test, $t(42) = 1.415, p = 0.164$. At post-test, the mean percentage for trained students (53.6%) was significantly higher than the mean percentage of control students (37.5%), $t(39) = 2.619, p = 0.012$. As shown in Figure 6, the training helped students gain a better ability to diagnose abnormal panoramic images as compared to control students.
Figure 6: Mean scores (percentage of correctly diagnosed abnormal images) of students and dentists at pre-test and post-test. Error bars represent standard error. $p < 0.05$
3.1.1.3 Incorrect Diagnosis of Abnormal Panoramic Images

The percentage of the incorrectly diagnosed abnormal panoramic images (identifying an abnormal image as being normal) was calculated.

The ANOVA of the mean percentage showed a main effect of time, $F(1, 70) = 10.177, p = 0.002$, indicating that overall performance at post-test was better than at pre-test. There was no main effect of groups or educational levels, $F(1, 70) = 1.846, p = 0.179$ and $F(1, 70) = 2.72, p = 0.104$, respectively. The time by groups interaction was significant, $F(1, 70) = 7.956, p = 0.006$.

However, the time by groups by educational levels interaction was not significant, $F(1, 70) = 0.686, p = 0.411$. The time and educational levels showed no significant interaction, $F(1, 70) = 0.001, p = 0.973$ nor did the groups and educational levels, $F(1, 70) = 0.336, p = 0.564$.

The results of the paired samples $t$-test, performed on the scores of control and trained groups collapsed across students and dentists, revealed a significance difference between the pre-test and post-test scores of trained groups, $t(37) = 4.813, p < 0.05$, but not for control groups, $t(35) = 0.18, p = 0.858$, indicating that performance of trained groups significantly improved from pre-test to post-test. There was a significant difference between the two groups at post-test, $t(72) = 2.885, p = 0.005$, but not at pre-test, $t(78) = 0.307, p = 0.76$. At post-test, the mean percentage of the incorrectly diagnosed panoramic images of trained groups (15.8%) was significantly less than of control groups (28.5%). As shown in Figure 7, the training decreased the percentage of incorrectly diagnosed abnormal panoramic images.
Figure 7: Mean scores (percentage of incorrectly diagnosed abnormal images) of students and dentists at pre-test and post-test. Error bars represent standard error. $p < 0.05$
3.1.2 Abnormality Detection Accuracy

The panoramic image contains both a complex upper portion, containing the orbit, skull base as well as nasal, maxillary and zygomatic bones, and a simpler image of the mandible. The complexity of the anatomical structure of the upper portion is due the greater amount of superimposition of hard and soft tissue structures in the region. Due to the differences in the complexity of these two parts of the panoramic image, the researcher thought to look into the abnormality detection accuracy in the complex part and in the mandibular part of the image, separately. Abnormality detection accuracy was determined by calculating scores for several outcomes. These scores were: (1) the percentage of correctly detected abnormal findings in the complex part, (2) the number of over-reported findings in the complex part, (3) the percentage of correctly detected abnormal findings in the mandibular part of the image, and (4) the number of over-reported findings in the mandibular part of the image. Separate scores were calculated for the pre-test and post-test. The effect of the systematic search strategy training was determined by measuring the difference between the pre-test and post-test scores.

3.1.2.1 Correct Detection of Abnormalities in the Complex Part of the Image

The ANOVA revealed that overall, participants’ performances were better at post-test than at pre-test as evident by a significant main effect of time, $F(1, 70) = 5.823, p = 0.018$. There was a significant main effect of groups, $F(1, 70) = 5.562, p = 0.021$, indicating that trained groups outperformed control groups. The main effect of educational levels was also significant, $F(1, 70) = 23.386, p < 0.05$, with dentists detected more abnormal findings in the complex part of the image than students. Neither the time by educational levels or the groups by educational levels interactions were significant, $F(1, 70) = 0.916, p = 0.342$, $F(1, 70) = 0.172, p = 0.68$, respectively. There was a significant time by groups interaction, $F(1, 70) = 4.39, p = 0.04$, as well a time by groups by educational level interaction, $F(1, 70) = 18.823, p < 0.05$. Subsequent 2-way mixed design ANOVAs, conducted separately on dentists’ and students’ scores, revealed that there was a significant time by groups interaction for students, $F(1, 39) = 27.479, p < 0.05$, but not for dentists, $F(1, 31) = 1.898, p = 0.178$.

Post-hoc analyses with paired samples $t$-tests was performed to compare students’ pre-test and post-test scores. The results revealed a significance difference between the pre-test and post-test
scores of trained students, $t(20) = -5.502, p < 0.05$, but not for control students, $t(19) = 1.823, p = 0.084$, indicating that trained students significantly improved in their abnormality detection ability from pre-test to post-test. To analyze the effect of groups, independent samples $t$-tests were performed to compare between the pre-tests scores of control and trained students and post-test scores. Results revealed a significant difference between the two groups at post-test, $t(39) = 3.965, p < .05$, but not at pre-test, $t(42) = -1.561, p = 0.126$. At post-test, the mean percentage of correctly detected findings, in this portion of the image, for trained students (41.67%) was significantly higher than the mean percentage for control students (11.25%). As shown in Figure 8, the training helped students gain a better ability to detect abnormal findings in the complex part of the panoramic image as compared to control students. The training, however, did not improve the performance of dentists.

![Figure 8: Mean scores (percentage of correctly detected findings in the complex part of the image) of students and dentists at pre-test and post-test. Error bars represent standard errors. $p < 0.05$](image)
3.1.2.2 Over-reported Findings in the Complex Part of the Image

The ANOVA of the mean number of over-reported findings revealed a main effect of time, $F(1, 70) = 19.839$, $p < 0.05$, indicating that overall, participants over-reported more findings on post-test than on pre-test. The number of the over-reported findings did not differ between students and dentists, as indicated by the lack of a significant main effect of educational level, $F(1, 70) = 3.312$, $p = 0.073$, time by educational levels, $F(1, 70) = 1.288$, $p = 0.26$, as well as a non-significant groups by educational levels, $F(1, 70) = 0.843$, $p = 0.362$, or time by groups by educational levels, $F(1, 70) = 1.673$, $p = 0.2$, interactions. There was a significant main effect of groups, $F(1, 70) = 12.722$, $p = 0.00$, and time by groups interaction, $F(1, 70) = 6.006$, $p = 0.017$.

A paired-samples $t$-test was conducted to evaluate the time effect on control and trained groups collapsed across students and dentists. The results showed that the number of over-reported findings significantly increased after training, $t(39) = 4.693$, $p < 0.05$. In contrast, there was no statistically significant difference between the number of the findings over-reported by control participants at pre-test and post-test, $t(35) = 1.155$, $p = 0.13$. Post-hoc independent samples $t$-tests revealed a significant difference between the two groups at post-test, $t(72) = 4.155$, $p < 0.05$, but not at pre-test, $t(78) = 1.743$, $p = 0.085$. The mean number of findings over-reported by trained participants at post-test (6.47) was significantly higher than the mean number reported by control participants (3.19). As shown in Figure 9, the training adversely increased the number of findings over-reported by students and dentists.
Figure 9: Mean scores (number of over-reported findings in the complex part of the image) of students and dentists at pre-test and post-test. Error bars represent standard error. $p < 0.05$
3.1.2.3 Correct Detection of Abnormalities in the Mandibular Part of the Image

Regarding the correct diagnosis of abnormal findings in the simpler image of the mandible, neither students nor dentists improved after training. The ANOVA of the mean percentage of correctly detected abnormal findings in the mandible revealed no main effect of time, $F(1, 70) = 2.36$, $p = 0.129$. The detection of abnormal findings in the mandible did not differ between control and trained groups, as indicated by lack of significant main effect of groups, $F(1, 70) = 0.69$, $p = 0.41$, time by groups, $F(1, 70) = 2.252$, $p = 0.138$, or time by groups by educational levels, $F(1, 70) = 0.683$, $p = 0.411$, interactions. However, the detection of abnormal findings in the mandible was affected by the educational levels, as indicated by a significant main effect of educational levels, $F(1, 70) = 22.15$, $p < 0.05$, and a significant time by educational levels interaction, $F(1, 70) = 5.843$, $p = 0.018$.

Paired samples $t$-tests revealed that dentists maintained their level of accuracy from pre-test to post-test, $t(32) = 0.529$, $p = 0.601$. In contrast, students’ performance dropped significantly over time, $t(40) = -3.303$, $p = 0.002$. In addition, independent samples $t$-tests were performed to compare the performance of students and dentists at pre-test and post-test. There was a statistically significant difference between students and dentists at both pre-test and post-test, $t(78) = -2.76$, $p = 0.007$, and $t(40) = -5.122$, $p < 0.05$, respectively. This indicates that the significant time by educational level interaction was a sole function of time effect on students. These results are presented in Figure 10.
Figure 10: Mean Scores (percentage of correctly detected abnormalities in the mandibular image) of students and dentists at pre-test and post-tests. Error bars represent standard error. $p < 0.05$
3.1.2.4 Over-reported Findings in the Mandibular Part of the Image

In contrast to over-reported findings in the complex part of the image, the training did not have an effect on the number of over-reported findings in the mandibular part of the image. This finding is supported by the lack of a significant main effect of groups, $F(1, 70) = 3.008, p = 0.087$, as well as non-significant time by groups, $F(1, 70) = 0.109, p = 0.742$, groups by educational levels, $F(1, 70) = 1.507, p = 0.224$, or time by groups by educational levels, $F(1, 70) = 0.03, p = 0.864$, interactions. There was, however, an overall decrease over time in the number of over-reported findings in this part of the image, as indicated by a significant main effect of time, $F(1, 70) = 15.748, p < 0.05$. As well, students over-reported more findings in the mandibular image than dentists, as indicated by a significant main effect of the educational levels, $F(1, 70) = 10.783, p = 0.002$. The time by educational levels interaction was not significant, $F(1, 70) = 0.032, p = 0.858$. These results are presented in Figure 11.

![Figure 11: Mean scores (number of over-reported findings in the mandibular part of the image) of students and dentists at pre-test and post-test. Error bars represent standard error. $p <0.05$](image)
3.1.3 Confidence Level

For each participant the average confidence level was calculated as a value from 0-10 for the following responses: (1) correctly diagnosed normal panoramic images, (2) correctly detected abnormal findings, (3) over-reported findings.

3.1.3.1 Confidence Level Regarding Correctly Diagnosed Normal Panoramic Images

The training had no effect on participants’ confidence level when they correctly diagnosed normal images. This was indicated by a non-significant main effect of groups, $F(1, 70) = 0.661$, $p = 0.419$. None of the time by groups, $F(1, 70) = 0.041$, $p = 0.839$, groups by educational levels, $F(1, 70) = 2.247$, $p = 0.138$, or time by groups by educational levels, $F(1, 70) = 0.009$, $p = 0.924$, interactions were significant. The ANOVA revealed that participants’ average confidence level decreased over time as evident by a significant main effect of time, $F(1, 70) = 4.806$, $p = 0.032$. The average confidence level of the correctly diagnosed normal panoramic images was also influenced by educational levels as indicated by a main effect of educational levels, $F(1, 70) = 21.344$, $p < 0.05$, with dentists being more confident than students. However, there was no time by educational levels interaction, $F(1,70) = 0.139$, $p = 0.71$. These results are presented in Figure 12.
Figure 12: Mean Average Confidence (correctly diagnosed normal panoramic images) of students and dentists at pre-test and post-test. Error bars represent standard error. $p < 0.05$
3.1.3.2 Confidence Level Regarding Correctly Detected Findings

The ANOVA of the confidence regarding correctly detected abnormal findings showed a main effect of time, \( F(1, 70) = 4.969, p = 0.02 \), no main effect of groups, \( F(1, 70) = 7.6, p = 0.386 \), and a significant time by groups interaction, \( F(1, 70) = 5.26, p = 0.025 \). The educational levels did not have an impact, \( F(1, 70) = 1.772, p = 0.188 \). Similarly, the time by educational levels interaction was not significant, \( F(1, 70) = 0.089, p = 0.766 \), neither was the groups by educational levels, \( F(1, 70) = 0.256, p = 0.614 \), nor was the time by groups by educational levels, \( F(1, 70) = 1.622, p = 0.207 \), interactions. Post-hoc paired samples \( t \)-tests revealed that confidence of trained participants in correctly detecting abnormal findings decreased over time, \( t(37) = -3.015, p = 0.005 \), whereas control participants had retained their confidence level from pre-test to post-test, \( t(37) = 0.193, p = 0.848 \). These results are presented in Figure 13.

![Figure 13: Mean Average Confidence (correctly detected abnormal findings) of students and dentists at pre-test and post-test. Error bars represent standard error. \( p < 0.05 \)](image)
3.1.3.3 Confidence Level Regarding Over-reported Findings

The confidence of calling the over-reported findings abnormal was not affected by any of the independent variables in the study. There was no main effect of time, $F(1, 70) = 0.361, p = 0.55$, groups, $F(1, 70) = 1.858, p = 0.177$, educational levels, $F(1, 70) = 0.434, p = 0.512$. Also, none of the time by groups, $F(1, 70) = 0.494, p = 0.484$, time by educational levels, $F(1, 70) = 2.167, p = 0.146$, groups by educational levels, $F(1, 70) = 0.202, p = 0.655$, or time by groups by educational levels, $F(1, 70) = 0.01, p = 0.921$, interactions were significant. These results are presented in Figure 14.

![Figure 14: Mean Average Confidence (over-reported findings) of students and dentists at pre-test and post-test. Error bars represent standard error. $p < 0.05$](image)
3.1.4 Search Time

In this study, search time (in seconds) represents time taken to determine whether the image was normal or abnormal. The search time averaged over all images in each test set was calculated. The ANOVA showed a main effect of time, $F(1, 70) = 16.583, p < 0.05$, indicating that overall search time was longer in post-test than pre-test. As well, the main effect of groups was significant, $F(1, 70) = 14.877, p < 0.05$, indicating that trained groups spent more time searching the panoramic images than control groups. However, the main effect of educational levels was not significant, $F(1, 70) = 0.049, p = 0.826$. The time by groups interaction was significant, $F(1, 70) = 32.313, p < 0.05$. The time by educational levels interaction was not significant, $F(1, 70) = 0.001, p = 0.806$, nor was the time by groups by educational levels, $F(1, 70) = 1.025, p = 0.315$.

Post-hoc paired-sample $t$-tests showed that there were statistically significant differences in the pre-test and post-test search time taken by trained participants, $t(39) = 44.342, p < 0.05$, and control participants, $t(35) = -7.611, p = 0.034$. The results indicated that for trained participants search time was significantly longer at post-test (80.8) than pre-test (36.45), while for control participants search time was shorter at post-test (33.88) than pre-test (41.49). As shown in Figure 15, the training significantly increased the time taken by students and dentists to search the image.
Figure 15: Mean time (sec) of students and dentists at pre-test and post-test. Error bars represent standard error. $p < 0.05$
3.2 Objective Two

The second objective was to determine whether training both novice and more experienced dentists in a systematic search strategy will improve the detection of multiple abnormal findings in panoramic images that contain multiple abnormalities. The comparison was made between the performance of the control and trained groups at post-test; where there were five panoramic images each contained two abnormalities.

3.2.1 Detection of Both Abnormal Findings in Panoramic Images with More than One Abnormalities

Each participant was given a score which represents the percentage of the panoramic images in which both abnormalities were detected. The means scores were compared by using a 2 X 2 (groups X educational levels) factorial analysis of variance to test the effects of the groups (control vs. trained) and educational levels (students vs. dentists) on the detection of both abnormal findings in the five panoramic images with two abnormalities at post-test set. Results indicated a significant main effect for groups, \( F(1,70) = 4.317, p = 0.041 \). Those who were given the training detected both abnormal findings in more panoramic images (\( m = 27.87\% \)) compared to those who had not (\( m = 17.00\% \)). There was also a significant main effect of educational levels, \( F(1,70) = 11.543, p = 0.001 \). Over all, dentists outperformed (\( m = 31.32\% \)) students (\( m = 13.55\% \)). There was, however, no significant interaction between the two factors, \( F(1,70) = 0.115, p = 0.735 \), indicating that the effect of training was the same for students and dentists. These results are shown in Figure 16.
Figure 16: Mean scores (percentage of images in which both abnormalities were detected) of students and dentists at pre-test and post-test. Error bars represent standard error. $p < 0.05$
3.3 Post-test Questions

When asked how easy or difficult the video instruction was to understand, 84.2% of the participants in the trained groups perceived that the video was easy to understand, 15.8% perceived it as neither easy nor difficult and none perceived it as difficult. In addition, when asked if the educational video and the checklist helped focus their search in areas that they might not have looked, participants responses were “yes” in 94.7%, in contrast to “no” in 5.3% of the responses and none as “cannot tell”.
4 Discussion

4.1 Discussion of Research Hypotheses

4.1.1 Research hypothesis 1.A

It was hypothesized that training dental students in a systematic search strategy would improve their radiologic interpretation performance with panoramic images. Providing a systematic search strategy training significantly improved the correct diagnosis of the abnormal panoramic images and the detection of abnormalities in the complex part of the image. The mean percentage of the correctly diagnosed abnormal panoramic images at post-test was significantly higher for trained students (53.6%) than for control students (37.5%). Also, the mean percentage of the correctly detected findings in complex part of the image was significantly higher for trained students (41.67%) than for control students (11.25%). These findings lend support to the effectiveness of a systematic search strategy for the analysis of complex radiographic images. This appears to contradict the findings of a previous study in the medical literature done with chest radiographs that found no significant effect of the use of a systematic search strategy upon students’ performance. In their study, Gale and Worthington\textsuperscript{23} found no statistically significant differences in the performance of the experimental and control groups. These authors attributed the lack of improvement to performing two tasks at once, searching the radiograph and adhering to a strategy, which may compete for attention allocation. It seems possible that the difference between our results and those of Gale and Worthington is due to the differences between the complexity and anatomic layout of chest and panoramic radiographs. Many more anatomic features per unit area need to be inspected in the complex part of the panoramic radiograph compared to a chest radiograph. Thus, with respect to this study, it is possible that the systematic search strategy training helped focus students’ attention on those anatomic features. The post-test question supports this assumption, in which 94.7% of trained participants responded ‘yes’ when asked if the video and checklist helped focus their search in areas that they might have overlooked.
Different results were obtained with the correct detection of abnormalities in the simpler mandibular part of the image. After training, there was no improvement in the ability of students to correctly detect mandibular abnormalities. Therefore, it is probable that training in a systematic search strategy does not result in better detection performance when the radiographic image is not complicated by overlapping anatomic structures. In this part of the image, there was a significant decrease in students’ performance over time. One possible explanation for this result is that post-test images, despite having been evaluated by dentists, may have been more difficult for students than the pre-test images.

Another interesting finding in this study was the significant increase in the number of the over-reported findings in the complex part of the image. This result is in agreement with those in the medical literature, in that the use of a systematic search strategy increased over-reported errors. In the current study, the over-reported findings were variations from normal anatomical features and image artifacts. This observation suggests that reporting a normal feature as abnormal might be the result of a lack of knowledge of the variations of the normal anatomy of the panoramic radiograph (pattern recognition error). For example, students in the present study were more likely to identify a submandibular gland fossa, the shadow of the soft tissue of the nose, and asymmetrical nasal concha as abnormal features. Figures 17 through 19 present example of these errors.
Figure 17: The shadow of the soft tissue of the nose (X₁) and the right submandibular fossa (X₂) were mistakenly identified as abnormal findings in this abnormal panoramic image. The true abnormality is in the maxillary left premolar-molar area (dotted arrows).

Figure 18: The incisive foramen (X₁) and the right nasal concha (X₂) were mistakenly identified as abnormal findings in otherwise normal panoramic image.
Figure 19: The superimposition of the cervical vertebrae over the mandibular midline was mistakenly identified as an abnormal finding in this normal panoramic image.
It is possible that the systematic search strategy training made participants search the image very carefully and this led them to question certain regions of the image that they would not normally investigate. This explanation is supported by the significant increase in the time taken by trained students to decide whether the image was normal or abnormal at post-test.

Alternatively, it is possible that the training altered students’ behavior and made them more cautious. Therefore, in an attempt not to miss any abnormality, normal variants were interpreted as abnormal features. Goldstein et al.\(^{38}\) studied the ability of dental students to decide whether a radiograph contained an abnormal finding including dental pathologies and anomalies, a processing defect or normal image. They found a high false positive errors rate (40.6%). The authors further reported that 75% of participating students indicated that they “would prefer to avoid missing an abnormality (as opposed to avoid false positive reporting) when confronted with a difficult interpretation”. Additionally, Mobley and Goldstein\(^{39}\) studied a similar effect by testing the effect of variable payoff conditions for students interpreting dental radiographs. The study revealed that the rate of false positive responses remained high even when payoff conditions penalized such responses.

The systematic search training provided in this experiment did not help students to correctly diagnose normal panoramic images. The control students started out with a clear advantage with a mean pre-test score of 67.14% as compared to the trained group (48.29%). This might not be reflective of any real difference in the knowledge between both groups but might instead be the result of chance. With this initial advantage, the mean post-test score of control students remained higher (60.0%) compared to the trained students (39.45%). It is possible that the training drew participants’ attention to normal anatomic structures that they inspected for the first time and therefore decided to identify them as abnormal. This explanation is supported by the significant increase in the over-reported findings made by trained students at post-test.

Without the use of an eye-position tracking device, it is not possible to determine the reason for missed abnormalities in this study. For instance, in a panoramic image with an abnormally widened inferior alveolar canal, such an abnormality could be missed if 1) the canal was not examined by the participants (search error), or 2) if it was examined by a participant following the systematic search instruction but failed to perceive this finding as abnormal due to lack of
knowledge about the features of the normal canal and what constitutes an abnormality (pattern recognition), or 3) if the participant perceived that the canal was not normal but decided that this could represent an image artifact or variation from normal (decision error). Research using the same images in the present study needs to be undertaken with the use of eye-tracking technology to investigate these proposed hypotheses.

To summarize, the results of this study suggest that systematic search strategy training is an effective approach in enhancing students’ ability to diagnose abnormal panoramic images and to detect abnormalities in the complex part of the image, but at the expense of an increase in the number of over-reported findings. In contrast, there was no statistically significant improvement in the detection of abnormalities in the simpler mandibular portion of the panoramic image. The training did not help students to recognize normal images. Thus, this study suggests some benefit in training students to use a systematic search strategy for the analysis of panoramic images. Further refinement is required so they do not lead to an increase in over-reported errors. These results have not been previously described in the dental literature.

4.1.2 Research hypothesis 1.B

It was hypothesized that training more experienced dentists in a systematic search strategy would not improve their radiologic interpretation performance because experienced dentists likely have already developed an effective search strategy. The findings of this study support our hypothesis and showed that the training did not help experienced dentists to correctly diagnose normal panoramic images or detect abnormalities in abnormal images. There was, however, a significant decrease in the percentage of incorrectly diagnosed abnormal panoramic images after training. This is the one finding that demonstrates an improvement in dentists’ performance after training. This might be the result of trained participants’ greater tendency to overcall panoramic images as abnormal and therefore might not be reflective of any real differences in dentists’ skill. Further research need to explore this finding to ensure it is not an artifact of our study.

There was, similarly to trained students, a statistically significant increase in the number of over-reported findings in the complex part of the image. Similar to students the over-reported errors were normal anatomical features and image artifacts. There was a significant increase in the time taken by trained dentists to search the images. Therefore, the same explanations proposed for the increase in the over-reported findings with trained students may explain this outcome in trained
dentists. It can be assumed that the search strategy made trained dentists search the image thoroughly and, consequently, brought certain parts of the image that they would not normally investigate to the forefront of their consciousness. Being in a study made them prefer to call those questionable areas abnormal rather than missing abnormalities. The consequences of over-reported errors in the interpretation of oral and maxillofacial imaging have not been discussed in the dental literatures. Over-reported findings may lead to over-treatment or additional diagnostic tests. In their examination of the causes of 182 errors in medical radiologic diagnosis, Renfrew et al. found that 15 errors were false positives. Of these 15 errors, 6 patients received surgery whereas others were subjected to further diagnostic tests. However, missing an abnormal finding could have a more serious patient outcome as it could result in delayed diagnosis and treatment.

In summary, this study found that the systematic search strategy training was not beneficial for experienced dentists. However, given the fact that the study was conducted with a convenience sample of dentists from one institution (KAU), caution must be applied in generalizing the findings of this study. Therefore, it is suggested to repeat the study with a different population of dentists who received different education and training.

4.1.3 Research hypothesis 2

The post-test set contained five panoramic images, each contained two abnormalities. It was hypothesized that training both novice and experienced observers in a systematic search strategy would improve the detection of both abnormal findings in panoramic images that contained two abnormalities. Participants who were given the systematic search strategy training detected both abnormal findings in more panoramic images (m = 27.87%) compared to those who were not given the training (m = 17.00%). However, because it is unknown how participants would have performed at baseline, we cannot reliably determine whether the training led to the observed differences in post-test.

Tuddenham speculated that satisfaction of search phenomenon is due to termination of search after detecting an abnormality. The results of Barbaum et al. do not support this explanation. In their study, eye-position analysis showed that observers continued to search images after reporting an initial abnormality and most missed lesions were inspected but then ignored. They, therefore, suggested that satisfaction of search could be attributed to errors in pattern recognition
or faulty decisions. In our study, it seems plausible that the training helped the participants to search the image thoroughly and/or to recognize the abnormal appearances of specific structures in the image. For example, the video emphasizes the importance of assessing the continuity of the floor of the maxillary sinuses as a discontinuity would signify an abnormality and of comparing the internal density of the maxillary sinuses. It also emphasizes examining the uniformity of the outer cortex and width of the inferior alveolar canal. However, because eye-tracking device has not been used, the actual mechanism that explains the outperformance of trained participants in detecting more than one abnormal finding remains unknown.

4.2 Study Limitations

Although this study has documented some benefits of training students in a systematic search strategy to analyze the panoramic images, which has not been previously reported in the dental literature, a number of limitations need to be noted.

First, a major weakness of the study lies in the fact that there was no empirical proof that participants applied the systematic search strategy since we were unable to track their eye movements across the images. Therefore, it is difficult to understand the mechanisms underlying the observed changes in the performance and the specific search strategy taught in this study. It was assumed that the inclusion of the checklist helped participants apply the instruction. Although 85% of trained participants indicated that the video and checklist helped to focus their search in areas of the image that they have not previously searched, there is no direct proof that the observers visually focused on each part of the image as outlined in the strategy. This could be substantiated with the use of eye-tracking technology. This device was not available to the researcher at the time of conducting the study. Conducting a study similar to the current one, with the inclusion of an eye-tracking device, would be useful.

A second issue that was not addressed in this study was subjects’ knowledge prior to enrolment in the study. To determine eligibility, the current study selected dentists with a minimum of five years of experience in the analysis of panoramic images, and used these images as part of their routine practice. This assumes that such an individual is more familiar with the normal panoramic features and the appearance of abnormalities. All students were recruited from the
same academic year and had completed the course work and the clinical training relevant to oral radiology interpretation. However, knowledge assessment was not done and their level of knowledge coming into the study was unknown. We hypothesize that lack of knowledge of variation in the appearance of normal anatomy and image artifacts could be a major factor, causing the increase in over-reported findings observed in the study. Experimentally, this hypothesis could be tested in a future study similar to this one, with the inclusion of training in the appearance of normal panoramic anatomy and its variations as well as image artifacts. If, as hypothesized, lack of knowledge is the cause of over-reported errors, then this future study should reinforce the diagnostic benefit of the systematic search strategy training by decreasing the over-reported errors rate.

Thirdly, adequate control over extraneous variables which may affect study outcome, such as fatigue and test time, was difficult to achieve in this study. Participants took the test as groups or individuals and test schedules were tailored to participants’ availability. For example, it is possible that participants who did the test at the end of the working day were more physically tired, fatigued and lacked energy than the ones who did the test earlier in the day. Also, it is possible that students who enrolled in the study earlier could have shared information and discussed their observations with their classmates who enrolled in the study later, and this may have affected the performance of the later students. The principal investigator attempted to minimize the influence of participants and especially students on one another by presenting the test images in a random order for each participant. Participants were also asked not to share any information with their colleagues until study completion.

This study is also limited by the fact that it was conducted in an artificial setting, which cannot simulate real practice setting. Participants were required to make a diagnosis without any clinical information. This, however, is often untrue in real clinical practice, since dental practitioners rarely work without the associated clinical information. Typically, dental practitioners are able to examine the patients and gather information about their medical and dental histories, and chief complaints. However, a possible disadvantage of having this information is that it may bias the clinician’s interpretation of the image. Also, not all jaw diseases have noticeable signs and symptoms. Therefore, the current study was specifically designed to assess the effectiveness of the systematic search strategy training without providing participants with clinical information that may influence the study outcome.
There were also limitations in the instrument used. Participants were not able to change their selection whether the image was normal or abnormal once a diagnosis had been made, nor were they able to delete a mark once placed. Comments made by participants during the tests indicated that this was an issue with the testing instrument. Also, the program lacks some features or tools for viewing digital images, such as image enhancement, contrast and brightness adjustment and a zoom function. These features are available in real-life practice and may enhance practitioner ability to detect abnormalities while examining a radiographic image.

The two sets of panoramic images were carefully chosen to be matched in difficulty and content in term of abnormalities location and features, such as widened canal, disruption in the continuity of the floor of the maxillary sinus, and condylar lesion. However, it was difficult to match other radiographic characteristics of the region of interest such as the size of the lesion, density, and effect on surrounding structure which could be more noticeable in a particular image than in its counterpart in the other set.

Although this study suggests some benefits in the use of systematic search strategy training for novices, one may argue that the improvement seen in students’ performance is due to the instructional design and presentation. The instruction used was a computer-based video training and included visual illustration of how to trace and analyze captured anatomical structures in a methodological manner. Based on this, it is possible that the results may be in part attributable to instruction module and not to the actual method of search strategy. For instance, would we observe the same improvement if the instruction was given as a conventional method of didactic teaching in a class-room? Further study is needed to determine what elements of the educational instruction were most beneficial to the students in order to design curricula and continuous education courses that most enhance students’ and dentists’ interpretation skills.

Due to the nature of the learning experience in this study and the fact that data were collected before and immediately after participation in the one training session, the long-term effectiveness of this single learning experience cannot be addressed. Also, this study examined the performance after viewing the training video for one time only and participants were not given the ability to view the video multiple times although some expressed the desire to do so. The study was designed in this manner in order to unify the viewing condition among participants. Based on information provided by participants, they found the training video and the instruction
very helpful and enlightening, but some did not feel that one viewing of the video was sufficient. Giving participants multiple opportunities to view the video would likely help them become familiar with the instruction so they can readily apply it while analysing the images.

Finally, this study was conducted with a convenience sample of students and dental practitioners from one university, King Abdulaziz University in Jeddah, Saudi Arabia. As a result, it would not be reasonable to assume that this population is representative of all students and dental practitioners from other institutions when curriculum might include more training, or continuing education courses are more available to dentists. Therefore, caution should be used in generalizing the results of this study.

4.3 Implications for Oral Radiology Education

This study suggests some benefits in the use of systematic search strategy training for novice observers. It also suggests some ways in which the systematic search strategy training can be ameliorated in order to ensure maximum efficiency of the training. In the current study, it was observed that variations in the appearance of normal anatomical features and image artifacts are frequent sources of over-reported errors. Therefore, further refinement with the inclusion of training in normal panoramic features and image artifacts is recommended as it may reduce over-reported errors. This could be achieved by increasing the number or length of time for focused radiographic interpretation sessions related to the identification of normal anatomy. Razmus et al.⁴⁰ found that the number of lecture hours provided and practice significantly affects students’ performance in the correct identification of normal anatomic features in panoramic radiograph. Students who received greater than three hours of lecture time on the panoramic image and had the opportunity to practice the technique on patients outperformed students who received 1 to 2 hours of instruction and did not have practice experience. To facilitate learning of the normal panoramic features, integration of the panoramic radiographic anatomy with gross skull anatomy may be beneficial for students. This may help students conceptualize how these anatomical structures are captured in the panoramic image, and integrate the anatomical appearances with radiologic interpretation.
Systematic search strategy training should not only tell the viewer to concentrate the visual attention to various anatomic structures in the image, but also should teach them the variable appearances of normal, and what constitutes abnormal appearances. For example, during training it is not only useful to tell the viewer to assess the cortical boundaries but it is equally important to mention that a discontinuity would signify an abnormality. This information may help the viewer to discriminate normal from abnormal appearances.

It would be interesting to explore if systematic search strategy training would be more effective when presented as an interactive program. For instance, would there be a difference in the performance if one novice group viewed the training video but had no practice and another group viewed the training video and immediately went through an interactive program analysing several panoramic images? This might substantiate the direct application of this strategy in the clinical setting. Clinical instructors could re-emphasize the application of this search strategy in analysis of the panoramic image and report writing sessions.

Lastly, the overall low detection rate of abnormalities, even by dentists, in this study is an indication that a significant number of abnormalities may be overlooked in clinical practice, with potential consequence on patient care. For instance, at baseline, control dentists incorrectly identified 46.8% of the abnormal panoramic images as normal. Trained dentists, at baseline, identified 53.6% of the abnormal images as normal. It is possible that dentists who prescribe and examine panoramic radiographs are not being given feedback on their performance, so they are not given the opportunity to learn from their practice. It would be interesting to find out if an interactive program that not only presents a search strategy but gives feedback via analysis of a series of practice panoramic images might improve performance. It would seem reasonable to encourage general practitioners and specialists to participate in this form of continuing education. When dental practitioners have better training in this important diagnostic procedure, patients are the ones who most benefit.

There is no evidence in support of the above mentioned suggested interventions. Therefore, more research is needed to maximize the efficiency of the systematic search strategy training.
4.4 Recommendations for Future Research

The current study will serve as a base for future research and provides important insight into the direction of research evaluating the visual search and interpretation of the dental radiographic images. In this experiment there was no proof that participants followed the systemic search strategy in examining the panoramic images. It is possible that the training caused the observers to be more cognisant of a series of anatomic structures within the panoramic image. Future research with the use of eye-tracking technology would determine whether participants applied the systematic search strategy while analysing the test images. This device may also highlight areas that are frequently missed during the visual examination of the panoramic image, and provide valuable data for educating students and dental practitioners. Most importantly, eye-tracking technology could be used to study the visual search behaviour of experienced oral radiologists, not only with the use of two-dimensional images such as the panoramic image but also with the use of more complex images, such as computerized tomography (CT) and magnetic resonance images (MRI). By exploring how experts search and allocate attention across different radiographic images, clinical instructors may be better able to train oral radiology residents, dentists and students. In the medical literature, considerable work has been done with the use of eye-tracking technology to study medical image perception and to understand how the radiologists interact with the visual information in the image to render a diagnostic decision.\textsuperscript{2, 41, 42} However, such studies are lacking in the dental field.

To further explore the reason for the increase in the over-reported errors rate reported by the trained group in this study, a subsequent similar study supplemented with the insertion of training in the appearance of normal variations of panoramic anatomy and panoramic image artifacts could be undertaken. It is possible that a combination of systematic search strategy training with normal anatomy training may decrease the over-reported errors. Also, if this study is replicated, one suggestion is to conduct knowledge assessment test prior to study participation to ensure that participants are at the same level of education and experience, as well as to provide participants with multiple training and practice sessions prior to the actual test run to ensure that they are familiar and could readily apply the instruction. The multiple training and practice sessions might increase the long-term retention of the search strategy. Future studies should also assess knowledge retention and long-term performance. Participants should be tested not only immediately after the instruction but also weeks later.
The hypothesis that clinicians examine radiographic images with clinical information because it may assist the viewer in detecting abnormalities provides the incentive to develop similar studies to investigate whether systematic search strategy training helps students and dental practitioners to overcome the influence of biases. This could be achieved by conducting a study in which participants are asked to diagnose a set of normal and abnormal panoramic images, without or with clinical information. Also as in the current study, participants could be divided into an experimental group, who would be given the systematic search strategy training and a control group, who would not be given instruction regarding the search strategy. However, some of the test images would be presented with a history suggestive of the correct diagnosis, some with a history suggestive of the incorrect diagnosis, and some without a history. The performance of the experimental and control participants would then be compared under the three different test conditions.

Because there was no statistically significant difference in the performance of the experienced dentists, conducting a future study with dentists graduated from different institutions when curriculum might include more training, or continuous education courses are more available would be worthwhile.

The current study focuses on the effectiveness of a systematic search strategy in the analysis of a commonly used, two-dimensional panoramic image. In recent years, cone beam computed tomography (CBCT) imaging has emerged and its uses and applications have expanded in the dental field.\textsuperscript{43} Due to its lower dose (in comparison with medical CT), dental practitioners from all specialities have come to use this technology in increasing numbers. However, there are concerns about whether dentists have adequate training and experience in the examination and interpretation of these complex images.\textsuperscript{44} CBCT imaging provides reformatted multi-planar slices and three-dimensional construction of the dental and associated maxillofacial structures. With such a large volume of data in a single acquisition, it is possible that the likelihood of errors in the interpretation of these images is more frequent than in a single two-dimensional image. Considering this, the examination and interpretation of the complete image data set and not only the region of interest must be undertaken, which is usually accomplished by scrolling through CBCT slices or three-dimensional images. Therefore, the goal of oral radiologists as experts and educators in this field is to develop a systematic search strategy to analyze these complex imaging techniques and to evaluate its effectiveness in future studies to allow dental practitioners
using this technology to render the most accurate and timely interpretation. This is a critical element for improving the quality of oral health care.
Chapter 5

5 Conclusions

The study set out to test the effectiveness of training students and more experienced dentists in a systematic search strategy for the analysis of panoramic images. Our results demonstrate that such training improves students’ performance in the diagnosis of abnormal panoramic images and detection of abnormalities in the complex part of the panoramic image but at the cost of an increase in over-reported findings which were mostly normal anatomic variants and image artifacts. No improvement was found in the detection of abnormalities in the mandibular part of the image. This study suggests some benefits in the use of systematic search strategy training for novices. Greater emphasis should be given to teaching anatomic variants on panoramic images, particularly in the more complex part of the image. Also, our results show that trained participants were more able to detect multiple abnormalities in an image than control participants. While this study did not demonstrate a statistically significant improvement in the performance of trained dentists, it would be worthwhile to repeat the study with a different population of dentists.
References


42. The role of perception in imaging: Past and future. Seminars in nuclear medicine Elsevier; 2011.


45. The American Dental Association Council on Scientific Affairs. The use of cone-beam computed tomography in dentistry: An advisory statement from the american dental
List of Appendices

Appendix 1

Students Recruitment Message

**Headline:** Participants needed for a study on the analysis of panoramic images

Do you have problems with the analysis of the panoramic radiograph? Do you feel that the panoramic radiograph is complex and uncertain how and where should you look?

I am conducting a study on the analysis of panoramic radiographs. I am looking for Participants from the 5th year dental students.

- **Who is conducting the study?**

The principle investigator is Dr. Hanadi Khalifa, a demonstrator in the Oral Radiology department at KAU and a master student at the University of Toronto, Canada, under the supervision of Dr. Michael Pharoah.

- **What is the study about?**

The analysis of panoramic radiographs is a complex task. This complexity is due to the superimposition of the hard and soft tissue structures of the region, variable distortion, ghost shadows, and low spatial resolution, which makes visualization of fine detail and subtle diseases difficult. Therefore, subtle abnormalities can be easily missed and this may jeopardize patient care and postponing necessary treatment. The goal of this study is to investigate observers’ performance with panoramic images with the use of specific instructions. The results from this study will have clinical significance by reducing errors and improving abnormality detection in panoramic radiographs. In addition, it will influence the design of Oral Radiology curriculum and continuing education courses.

- **How much time will it take?**

If you agree to participate, you are required to participate in two sessions, one week apart. Each session will take approximately 30 minutes.
• What am I going to do?

In each session, you will view 15 panoramic radiographs using a computer program, some of which contain abnormalities and some have no abnormalities. You will be asked to determine if they are normal or abnormal, mark the location of the abnormalities directly on the panoramic images and finally assign a confidence level for each finding.

• What are the possible risks?

There are no risks to you from participating in this study. No individual information will be disclosed. You will be assigned an identification number (ID) to allow me to match your answers in the two sessions. Your identity can never be revealed since no one can identify these ID numbers. I would like also to emphasize that the study does not aim to assess your academic or professional abilities.

• What are the benefits?

A possible benefit will be the improvement in your ability to analyze the complex panoramic radiograph. After completion of the study, the researcher will devote time to review all the images included in the tests and to answer any questions.

This experiment will be done at the Faculty of dentistry at KAU in April, 2012. Exact dates are to be determined. The study time will be scheduled according to your convenient and availability. If you are interested in participating, or would like to get more information, please contact dr.hkhalifa@gmail.com.

I would very much appreciate your participation

Hanadi Khalifa
Appendix 2

Experienced Dentists Recruitment Sheet

**Headline:** Participants needed for a study on the analysis of panoramic images.

We are conducting a study on observers’ performance with panoramic images. We are looking for dentists or dental specialists, who routinely view panoramic images in their practice and have a minimum working experience of 5 years, to participate.

If you agree to participate in this study, you are required to participate in two sessions, one week apart. Each session will take approximately 30 minutes. In each session, you will view a series of 15 panoramic images, and you will be asked to determine if they are normal or abnormal, determine the location of the abnormality and assign a confidence level.

After completion of the study, the researcher will devote time to review all the images included in the tests and answer any questions. This experiment is located at the Faculty of dentistry at King Abdul-Aziz University. If you are interested in participating in this study, or would like to get more information, please contact dr.hkhalifa@gmail.com.
Appendix 3

Panoramic Images Evaluation Sheets

<table>
<thead>
<tr>
<th>Agreement Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Normal</th>
<th>Abnormal</th>
<th># of Abnormal Findings</th>
<th>Difficulty level from 0-10</th>
<th>Obvious</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 4

The Checklist

1. The periphery and the corners of the panoramic image.
The following structures, when visible, are assessed:
   - The orbits.
   - The skull base, including the articular process of the temporal bone.
   - The contents of the neck, including the cervical spine, styloid processes, pharynx, submandibular gland regions and hyoid bone.

2. The outer cortices of the mandible (starts from the base of the anterior surface of the ramus on the right then trace the cortical boundaries to the anterior border of the ramus on the left side).

3. The outer cortices of the maxillary sinuses (start from the most superior aspect of the posterior walls, floors and then the medial walls).

4. The zygomatic processes of the maxillae, the pterygomaxillary fissures and the zygomatic bones.

5. The internal density of the sinuses (compare the right and the left sides).

6. The structures of the nasal cavity (the nasal floor, concha and nasal septum).

7. The bone pattern of the mandible.

8. The bone pattern of the maxilla.

9. The alveolar processes and teeth (start from the posterior of the first quadrant to the posterior of the fourth quadrant in a clockwise direction)
Appendix 5

Written Instructions

(For all participants at pre-test and for the control groups at post-test)

You will be shown a series of 15 panoramic images. Some will be completely normal, and others will contain one or more abnormal findings. We would like you to consider that each panoramic image belongs to one of your patients, and your task is to view the images as you typically would, and spend the usual amount of time on each image. When you are satisfied with the search, you will be asked to:

1. Determine whether the panoramic image is normal or contains abnormal radiographic finding(s). The abnormal radiographic findings are the findings that are not consistent with normal anatomy as seen in the panoramic image and you think it might require consultation, further investigation or treatment. Although the dentition should be viewed for the presence of abnormalities, caries, periodontal bone loss, and impacted and missing teeth are not considered abnormal findings.

2. Mark the location of the abnormal finding(s), if present, on the panoramic image. The location of the mark should be as close to the centre of each finding as possible.

3. Mark your level of confidence for each finding on a scale between “suspicious” and “definite”.

There is no time limit for viewing.

Technical instruction: The marking on the panoramic image and the confidence scale will be done by left clicking on the area of interest. If you have made a mistake, place your mouse over the “X” mark and left-click again to remove it. The program will not allow you to go back to the previous slides once you had continued on.
Appendix 6

Written Instructions

(For the trained groups at post-test)

You will be shown a video that demonstrates a systematic approach to analyze the panoramic image. The video has a player control that permits you to pause if you wish to take notes. The video will be played once and you have 20 minutes to complete it. After the video, you will be shown a series of 15 panoramic images. Some will be completely normal, and others will contain one or more abnormalities. For each panoramic image, we would like you to follow the video instruction. We will also provide you with a checklist, which is a summary of the video, to guide you in viewing each section of the panoramic image. After viewing each panoramic image, you will be asked to:

1. Determine whether the panoramic image is normal or contains abnormal radiographic finding(s). The abnormal radiographic findings are the findings that are not consistent with normal anatomy as seen in the panoramic image and you think it might require consultation, further investigation or treatment. Although the dentition should be viewed for the presence of abnormalities, caries, periodontal bone loss, and impacted and missing teeth are not considered abnormal findings.

2. Mark the location of the abnormal finding(s), if present, on the panoramic image. The location of the mark should be as close to the center of each finding as possible.

3. Mark your level of confidence for each finding on a scale between “suspicious” and “definite”. At the end of the study you will be asked to fill out a brief questionnaire.

There is no time limit for viewing.

Technical instruction: The marking on the panoramic image and the confidence scale will be done by left clicking on the area of interest. If you have made a mistake, place your mouse over the “X” mark and left-click again to remove it. The program will not allow you to go back to the previous slides once you had continued on.
Appendix 7

Consent Agreement

Introduction
The goal of the current study is to investigate observers’ performance with panoramic images. The results from this study will have clinical significance by improving performance in the analysis of the panoramic image and will influence the design of undergraduate and postgraduate Oral Radiology curricula and continuing education courses.

Study Procedure
If you agree to participate in this study, you are required to participate in two sessions, one week apart, for a total time commitment of approximately one hour. In each session, you will view a series of 15 panoramic images, some of which contain abnormalities and some have no abnormalities. You will be asked to view each image and decide if it is normal or contains abnormal findings.

Benefits and Risks
There are no known risks to you from participating in this study. The results do not reflect your academic or professional abilities. A possible benefit will be improvement in your ability to analyze the panoramic image.

Subject Rights
Your participation in this study is voluntary. You may withdraw from the study at any time without penalty. You are encouraged to participate in the two sessions, but you are under no pressure from staff members or anyone to do so. If you are a student, your refusal or withdrawal from the study will not affect your academic evaluation or marks by any means. You will be provided with an email address if you need to ask questions about the study at any time.

Confidentiality
All information collected about you in this study will be confidential. No individual information will be disclosed. You will be assigned an identification number (ID) to allow us to match your answers in the two sessions. However, this information is confidential since no one can identify these ID numbers except for the principal investigator and thus your identity can never be revealed. The data will be securely stored in password-coded computers. No one will have access to these data except the principal investigator. Five years after completion of the study, all information will be destroyed. Researchers will use the results of this study to write scientific
papers and present at scientific conferences. Participants name or personal details will not appear in any report or presentation of this research.

**Contact**

Dr. Hanadi Khalifa is the principal investigator under the supervision of Dr. Michael Pharoah (michael.pharoah@dentistry.utoronto.ca). If you need any further information about the study, you can contact Dr. Hanadi Khalifa by email at (dr.hkhalifa@gmail.com).

**Consent Agreement**

I acknowledge that the procedures of this study have been explained to me clearly. I had the opportunity to ask questions, and any questions were answered to my satisfaction. I am aware that I may ask further questions at any point. I have been provided with contact information for the research supervisor of this study. I am aware that my participation is voluntarily. I can withdraw from the study at any time. In addition, my participation or withdrawal will not affect my academic evaluation.

- [ ] I agree to participate
- [ ] Disagree to participate

Name (please print): ________________________

Signature:__________________________ Date:________________
Appendix 8

Trained Groups Post-test Questions

1. How easy or difficult was the video instruction to understand?
   - Easy
   - Difficult
   - Neither easy nor difficult

2. Did the video and the checklist help focus your search in areas that you might not have looked?
   - Yes
   - No
   - Not sure
Appendix 9

Ethics Approval by the Research Ethics Committee of King Abdulaziz University

To: Hanadi Mohammed Khalifa
Department: University of Toronto/Craniofacial Radiology
Subject: REC Decision
Proposal No.: 003-11

The committee has reviewed your proposal entitled: The Effectiveness of A Systematic Search Strategy in the Analysis of Panoramic Images.

Please be advised that with respect to: (1) the rights and welfare of the individual(s) involved, (2) the appropriateness of the methods to be used to secure informed consent, and (3) the risks and potential benefits of the investigation, the Committee considers your project:

☑ Exempt
☐ Acceptable
☐ Acceptable with reservation noted (see attached letter)
☐ Not acceptable for reasons noted (see attached letter)
☐ Follow-up: The Committee wishes to have a status report on this project on

SIGNED FOR THE COMMITTEE BY:

Ali AlGhamdi BDS, MS, FRCD(C)
Vice Dean for Postgraduate Studies & Scientific Research
Director of Research Ethics Committee
Faculty of Dentistry, King Abdul Aziz University

Approval Date: 11/8/15

Type Project: ☑ NFW ☑ RENEWAL
Human Risk: ☑ Yes ☑ No
Source of Support: ☑ Outside Funding ☑ Department/other
Agency (Potential): ☑ Agency No.

Are any of the following involved: ☑ Yes ☑ No

If yes, which category (ies):

☐ Minors ☐ Fetuses ☐ Abortuses ☐ Pregnant Women
☐ Prisoners ☐ Mentally Retarded ☐ Mentally Disabled

CC: Dean
Department Chairman
Follow-up (Book)
File

Tel: 6403443
Fax: 6403318
P.O. Box 80209 Jeddah 21589
E-Mail: den.Faculty@kau.edu.sa
den.Faculty@kau.edu.sa
Appendix 10

Ethics Approval by the Research Ethics Board of the University of Toronto

PROTOCOL REFERENCE # 27081

December 13, 2011

Dr. Michael Pharoah
FACULTY OF DENTISTRY

Dr. Hanadi Mohammed Khalifa
FACULTY OF DENTISTRY

Dear Dr. Pharoah and Dr. Hanadi Mohammed Khalifa,

Re: Your research protocol entitled, "The effectiveness of a systematic search strategy in the analysis of panoramic images"

ETHICS APPROVAL

Original Approval Date: December 13, 2011
Expiry Date: December 12, 2012
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 Gwewu
REB Manager

UNIVERSITY OF TORONTO
OFFICE OF THE VICE PRESIDENT, RESEARCH

OFFICE OF RESEARCH ETHICS
McMurtry Building, 12 Queen's Park Crescent West, 2nd Floor, Toronto, ON M5S 1S8 Canada
Tel: +1 416 946-3271 • Fax: +1 416 946-5795 • ethics.review@utoronto.ca • http://www.research.utoronto.ca/for-researchers/administrators/ethics/