Designing and Using Simulation to Study Expert-Novice Differences in Correlating Medical Imaging with the Physical Exam

By

Lonie R. Salkowski

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Curriculum and Instruction)

at the

UNIVERSITY OF WISCONSIN-MADISON

2017

Date of final oral examination: 4/21/2017

The dissertation is approved by the following members of the Final Oral Committee:

Rosemary Russ, Assistant Professor, Curriculum and Instruction
Richard Halverson, Professor, Education Leadership Policy Analysis
Kurt Squire, Professor, University of California – Irvine, Department of Informatics
Matthew Berland, Assistant Professor, Digital Media in Department of Curriculum and Instruction
Leema Berland, Associate Professor, Curriculum and Instruction
Rock Mackie, Emeritus Professor, Medical Physics; Human Oncology; Biomedical Engineering

DESIGNING AND USING SIMULATION TO STUDY EXPERT-NOVICE DIFFERENCES IN CORRELATING MEDICAL IMAGING WITH THE PHYSICAL EXAM

Lonie R. Salkowski

Under the supervision of Assistant Professor Rosemary Russ

at the University of Wisconsin-Madison

The ability to correlate anatomical knowledge and medical imaging is foundational to radiology and to medicine more generally. For example, correlating the physical exam findings of the thorax and the chest x-ray findings can make the difference whether a patient is treated with antibiotics or supportive care with rest and fluids. Experts do this correlation well, but we have little understanding of how they learn to do so. Knowing where our novices' understanding begins in making this correlation can aid us in helping them transition to being experts.

Studying how experts and novices make these correlations is challenging because there isn't an efficient and direct way to make or observe how this task is performed. However, simulation in medical education has been shown to provide opportunities to learn about clinical conditions/treatments and acquire new skills in a safe protected environment that otherwise would not be possible within the context of daily patient care. I developed a simulation tool that incorporates medical imaging with the physical examination to directly investigate how experts and novices reason about this correlation. Not only did the simulation allow the study of the thinking involved in correlation, but also medical students increased their confidence in making this correlation by using this simulation tool.

In this work I used this simulation tool to explore how novices and experts make the correlation between the anatomy of the physical examination and cross sectional imaging. Novices were fourth year medical students and experts were radiologists. A clinical interview was developed to gain access to thought processes and establish the context of these thoughts during specific tasks. This interview allowed flexibility for exploration of the participant's thoughts but also a framework to compare between and across participant groups. The participants were encouraged to verbalize their thoughts while performing various tasks with the simulation tool. These tasks included localization of a target image within the body, assessing the correctness of a peer localization of a target image and the normality of an anatomical structure, and evaluation whether a biopsy approach was safe and any possible alternatives to the approach. The data collection used mixed methods, allowing me to keep my research question central and use the particular method (qualitative or quantitative) best suited to that investigation. Data such as time, probe movements and imaging plane were collected within the simulation tool software. Clinical interviews were transcribed and the data was coded using action codes modified from Crowley, Naus, Stewart and Friedman (2003). The data was analyzed using one and two–way ANOVAs, independent t-tests and Chi-square analysis using IBM SPSS Statistics, version 23.

When participants were asked to localize radiological images on the physical body, experts and novices demonstrate differences consistent with findings from other expert-novice studies. Specifically in this study experts were faster than novices (simulation localization time), and used fewer movements to get to target when matching the final image to the target image (simulation probe movements). Experts approached each event

from a clinical perspective and drew upon past experience, where as novices evaluated each event in literal context (qualitative action codes). Novices often used non-anatomic cues and incorrect terminology (qualitative action codes), and experts recognized meaningful patterns not noticed by novices (qualitative action codes).

Although experts and novices differed in their ability to locate target images, they demonstrate similar skills in recognizing correctness in peer assessments of correlating medical imaging with the physical body. Both novices and experts were able to recognize errors of peers related to incorrect probe localization for a target image, peer assessment of normality of the anatomic structure and whether the biopsy approach was safe. Experts were able to offer a deeper and more complete explanation to their evaluation of the peer assessments and reasoning for alternate biopsy approaches. I attribute these differences to differences in recall and recognition. Recognition is easier than recall because it bypasses the retrieval process, something that the experts have developed with experience and knowledge acquisition.

This study of cognitive differences in novices and experts may provide meaningful insights to develop educational programs that promote learning. Currently, although both are important for making medical decisions and diagnoses, human anatomy associated with the physical exam and imaging studies are often taught separately and disconnectedly. This work indicates the importance of teaching them together. Further, it suggests that novices do bring prior knowledge and experience to a new learning situation, and novices are able to recognize concepts within a domain. These two factors should be capitalized upon when developing new curriculum.

Table of Contents

List of Figures

Chapter 2

Chapter 3

Chapter 4

List of Tables

Chapter 2

Chapter 3

Chapter 4

Presentations

1. 'Experts and novices engage in different cognitive processing when correlating anatomy and imaging', accepted for oral presentation at Association of University Radiologists (AUR) 65th Annual Meeting, May 10,2017, Hollywood, FL.

- Content material from chapter 3
- Abstracts do get published

2. 'Cognitive processing differences of experts and novices when correlating anatomy and cross sectional imaging', accepted for oral presentation at Medical Image Perception Society (MIPS) XVII Biennial Meeting, July 12-14, 2017, Houston, TX.

- Abstracts are not published
- Organization is aware of AUR meeting presentation

3. 'Novice-expert differences and similarities in correlating and evaluating medical images within a simulation setting', submitted for oral presentation at Radiological Society of North America (RSNA) 103rd Scientific Assembly and Annual Meeting, November 26-December 1, 2017, Chicago, IL

- Content combines findings from chapter 3 and 4
- Will be notified of abstract results late July, 2017
- Organization is aware of AUR and MIPS meeting presentations

Acknowledgements

My reasons and aspirations for attaining an advanced degree in education transformed during my educational process. Much of this was due to the influence of amazing mentors along the way. I am blessed to have the range of support from the educational community coming into this field after already attaining a career. I could have been content to finish my career in medicine, as most would. However, I wanted to explore the field of education in more depth to make a larger impact on the field of medical education, and more specifically within radiology.

The process of getting a PhD is unique for each person. Beyond the core requirements, there are so many possibilities you can take to develop your niche. I am thankful that my mentors have supported my interests and allowed me to take those chances.

Rich Halverson was the first to encourage me to pursue a PhD. Prior to that I thought I was content with my Masters in Education Leadership and Policy Analysis. He saw something that I wasn't listening to or aware of in the possibilities that a doctoral degree in education could afford.

Kurt Squire, for taking me on as an intern within Games Learning and Society (GLS). I learned so much about how games can be educational and the importance of thinking about a story line to incorporate into games for them to be meaningful and played. Kurt also took the chance on me of accepting me as his doctoral student. Navigating the higher education scene is much more complex today than when I went through the first time. I am grateful for this opportunity. I hope our paths cross again in the future as I think there are

yet great things to be developed with medical education games and understanding how learners simulate the information into meaningful context.

Rosemary Russ for agreeing to advise me in the last years of my PhD. I learned so much from Rosemary. I appreciate her view of education and her continual push to help me develop and nurture my ideas. She was the one who broke the news to me that I am a cognitivist. I thought I was a pure constructivist, but as she would say – you are a cognitivist constructivist. I cannot thank Rosemary enough for being there for the difficult steps of fostering my data into meaningful works that I can share with those beyond the university. She helped push me to understand concepts beyond what I knew I was capable of doing. In this process, although she has been my advisor, I have also come to value her as a friend and colleague. I know we will continue to work together bridging the disciplines of education and medicine.

Rock Mackie, I am thankful for his continued support over the years. He has pushed me to see my work in a more holistic view and with the encouragement to look at the role of my interests and how they can fit into a business model and venture. Thank you for the willingness and importance of serving on my committee. It was so helpful to have your viewpoint from the medical world to help explain difficult concepts during my defense.

Mathew and Leema Berland, I am thankful for you being on my mentor committee and offering me candid and sometime difficult questions. Those definitely will prepare me for taking my material outside of the University.

Cole Cook, for patiently guiding me through the process of statistical analysis, and for helping me to understand the specific tests and the results that they delivered. Your help, made the analysis and writing of my dissertation a whole lot easier.

My friends and colleagues, especially Ed Bersu and Beth Meyerand, whose unceasing support on good and bad days has been a strong-hold for me to keep forging ahead. Thank you for all your encouragement, I am forever grateful. I am also thankful to my clinical colleagues who really have supported and encouraged me, especially during this tough last year of writing and defending. I appreciate the adjustments in the clinical schedule when I had looming deadlines. It truly relieved some of that extra stress.

Last, but definitely not least, is my family. My family has been my foundation in all I have been able to accomplish. Their support and prayers have not gone unnoticed. I have been truly blessed to have a supportive and loving family, which includes my parents, Ted and Kathy Salkowski, and my siblings Lynn, David, Lori and Lea. It was such an honor to have them all present at my graduation ceremony and celebration. That feeling is going to be hard to beat. I love you.

Chapter 1

Introduction

Imagine two medical students one that was taught human anatomy only through the study of the physical exam. The student will develop their skills with expertise as they learn to palpate normal and abnormal organs, and auscultate the various sounds of health and disease in the body. Then imagine a second medical student whose only exposure to human anatomy is through looking at medical imaging studies. Over time this student will develop the ability to differentiate the subtle abnormalities that discriminate the normal and abnormal imaging appearance of organs and disease.

Now these two students meet on clinical service. They have very different views of medicine, however neither of their knowledge of anatomy is complete. As these two navigate the clinical service, the one who knows the anatomy through physical exam realizes that if they just had the imaging knowledge, that their physical examination skill could be even better. And the student with the imaging knowledge realizes that the application of the physical exam skills to their imaging knowledge could open up a new approach to understanding the anatomy in medical images.

Although this example is extreme, it does reveal one of the issues that I have observed in teaching medical students over the last eighteen years. Much of anatomical medical education has been taught in silos similar to these two students. Anatomy has been taught as an independent course, with some introduction of medical concepts and imaging. But the enduring clinical segment of human anatomy, the medical imaging, has been treated as a periphery rather than as a practical extension of anatomy. Likewise, the physical examination skills are taught in a separate course that often does not align with

the anatomy that the students are learning. Sometimes they examine areas of the body without adequate understanding of the underlying anatomy. In these instances, there is a crash course application of the 'relevant' anatomy.

This hypothetical example conveys the framework for my thesis project, with the purpose of my thesis being to investigate the proficiency of correlating the anatomic physical examination with medical imaging studies. This will be investigated using the novice-expert paradigm to provide comparisons at the extremes of the knowledge spectrum. The implication for investigating this correlation is to understand if there is a need to integrate this type of correlation instruction in the medical school curriculum.

The remainder of this introduction will provide the background context of this project, beginning with describing the components of a traditional medical education. Then, describe the major research paradigms in medical education. Lastly, I will describe where this thesis project attempts to fill a gap in this research.

Field of Medical Education

Traditional medical education – course based

The four years of medical school are referred to as undergraduate medical education, and residency training as graduate medical education. The first two years of undergraduate medical education are referred to as pre-clinical years and structured to deliver the "basic sciences" of medicine. Students take courses in human anatomy (gross anatomy, neuroanatomy, embryology and histology), physiology, genetics, biochemistry, psychology, pharmacology, and pathology. The second two years are clinical based rotations that range from internal medicine, obstetrics-gynecology to surgical specialties.

This separation of courses is the most efficient and cost effective way to disseminate this large amount of content in the preclinical years however this method also contributes to the lack of correlation between disciplines

Due to the continued pressure to see patients and provide institutional income generated from seeing patients, there has been an offset of the basic science content to 'basic science professors'. This component of instruction has few practicing physicians actively involved in teaching. This does place a strain on knowledge content, suffering from over-detail, which is not necessary for most physicians to practice medicine, and lack of emphasis of clinically important concepts. Also the terminology used in the basic sciences, especially in anatomy, often differs from that of clinical medicine, thus causing student confusion and the need to 're-learn' material (Kachlik, Baca, Bozdechova, Cech, & Musil, 2008; Rector, 1999).

Because content is taught in courses, there is a lack of scaffolding and spiraling as coined by Bruner. Bowen (2006) noted that in clinical settings, student's recall of basic science knowledge from the classroom was often slow, awkward or absent. Students tend to organize their knowledge according to the structure of the curriculum. Due to the lack of connections made between their knowledge and the specific clinical encounters, the material needs to be readdressed to draw upon stored memory.

During the clinical training years of medical school, the students rotate through various clinical experiences. There is an immersion into clinical medicine and learning about the field of medicine with the application of the core knowledge from the first two years. Medical students spend a limited amount of time in radiology during their education. There has been a larger emphasis on the primary care fields. Many schools

don't have required radiology rotations, thus leaving radiology and its importance in healthcare taught by non-radiologists. Even on radiology rotations, the students have required didactic content and spend little time interacting with imaging studies or radiologists in the reading room.

The traditional medical school curriculum does not specifically address the correlation of the physical exam, medical anatomy and medical imaging studies. More or less it is implied that this correlation will be made within the clinical setting or by working through a problem case. Thus this leads to a discussion of how anatomy and the physical exam skills are commonly taught in medical school.

Teaching of anatomy and physical exam

Learning medical anatomy is a time intensive course, occupying over 100 hours of didactic content and laboratory experience in gross anatomical dissection (Albanese, 2010). Beyond the study of morphology or the structure and function of the body, anatomy is also the gateway into learning the geography of the biological entity (Guttmann, Drake, & Trelease, 2004). Experiencing anatomical human dissection involves the hands-on and personal or emotional aspects of the educational journey to become a physician (Granger, 2004).

Due to outside pressures for students to spend less time in the classroom, there has been drastic cuts in student contact hours, thus all course content needed to be cut – one course hardest hit was anatomy. Most schools no longer offer dissection due to these cuts and lack of professors to teach this material.

Some institutions have replaced or supplemented gross anatomy courses with innovative electronic teaching and learning tools (Jenkinson, 2009). These innovations have lead to great discussion about how anatomy should be taught. Some of these changes have been the result of curriculum reform and others due to financial reasons. Still others have suggested that although dissection is interactive, there is a need for more experiential learning in medical anatomy and propose a combination of various educational methods: dissection/prosection, interactive multimedia, procedural anatomy, surface and clinical anatomy (Asad & Nasir, 2015; Azer, 2013; Bergman et al., 2012; Dusseau, Knutson & Way, 2008), simulation (Hyde, Erolin & Ker, 2012) and medical imaging (Drake, 2007; Gunderman & Wilson, 2005; Jones, Olafson, & Sutin, 1978; Rizzolo, et al., 2010; Sugand, Abrahams, & Khurana, 2010; Wilson & Nava, 2010). All of these attempts have been to make the anatomy more relevant to the caring of patients.

Most physical examination skills are learned through an apprenticeship practice (Verghese, Brady, Kapur, & Horowitz, 2011; Wenrich, et al., 2011), and small group courses with standardized patients. In the preclinical years, students often attend clinic once or twice a month with a preceptor. During that time, the student sees patients with the preceptor. At set points in the curriculum, the students are also assessed on their ability to perform elements of the history and physical, and a complete examination. The goal is for students to be able to perform a basic physical exam prior to the clinical rotations.

These two components of medical education – anatomy and physical examination skills are taught separately. Anatomical education is variable, and even when dissection is a component of the course; the clinical skills are not often incorporated with the dissection to reinforce the anatomy. Furthermore, once the dissection has begun the reference to the

overlying physical landmarks are lost. Additionally, there is limited or lack of correlation of this anatomy with imaging studies.

This correlation or lack of correlation between the physical exam and medical imaging has not been studied within medical education. The research in medical education has largely focused on outcomes and based more around policy rather than what knowledge or reasoning that individual students possess. Thus emphasizing the importance for this investigation that explores the cognitive processes involved in understanding the correlation between medical imaging and the physical examination by novices and experts. To further explain where current research in medical has focused, I will briefly discuss these in relation to medical education, anatomy and medical imaging.

Trends in medical education research

Three bodies of research need to be explored: medical education research, medical anatomy research and research related to medical imaging. Within these domains, the major components of research as pertaining to this thesis will be briefly described.

Medical education research

Research related to medical education has largely centered on curriculum (Lee, Whelan, Tannery, Kanter, & Peters, 2103; Regehr, 2004), learning environments (lecture-based, problem based, team-based, and active versus passive learning strategies) (Lee, et al., 2103; Schaefer, et al., 2011), assessments, core competencies (Albanese, Mejicano, Mullan, Kokotailo, & Gruppen, 2008), professionalism, and technology that assists learning

(Davis, Crabb, Rogers, Zamora, & Kahn, 2008; Lee, et al., 2103; Schaefer, et al., 2011). The areas of research are often based around outcomes – something that can be used to compare one class year to the next, or to national standards (Chen, Baucher, & Burstin, 2004; Dauphinee & Wood-Dauphinee, 2004). These are student level outcomes, with no way to assess if these learning or curricular differences have an effect on eventual patient care (Collins, 2006). Research that is more directly related to student characteristics, have been interested in student learning styles and motivation, and diversity within medical schools (Mahant, Jovcevska, & Wadhwa, 2012).

There have been two large limitations of research in medical education. One, there has been a lack of productive research programs that are able to address and advance the 'big questions' in medical education (Regehr, 2004). Secondly, there have been limited funding opportunities in medical education and most projects are conducted on a small scale without any funding. These studies add to the body of work within medical education, but still do not impact the 'big questions'.

Research in anatomy education

Research within the anatomy courses in medical education has been performed in the attempt to make the experience more clinically relevant. Kotze and colleagues (2012; 2013) developed the concept of the 'translucent cadaver'. They performed full body digital x-rays of the cadavers to be used for dissection. Full sized images were printed out and posted on the wall near the respective dissection tank. The students could draw on these images and compare what they saw on the 2-dimensional image to the structures at dissection. Secondly, two similar studies performed by McLachlan and DeBere (2004) and

Patten (2007) projected anatomic images, images from the visible human project or radiographs on student volunteers. These images were projected on the skin surface or tight fitting clothing allowing the appropriate spatial relationship of the image and the living anatomy. All three of these studies involve exploring the surface anatomy as it relates to the 2D image. There is a lack of correlation of the cross sectional anatomy and the exploration of how these images relate back to the physical examination. Additional research has been performed with the majority of this research directed by anatomists. Their deficiency in real clinical skills, has limited this direct correlation.

Medical imaging research

Radiology is a complex skill that involves the ability to combine information from visual pattern recognition, anatomical knowledge, knowledge of pathological processes, and patient-specific information (Rubin, 1989). Research in medical imaging has been based on technological developments (imaging equipment, viewing devices and technology transfer), residency training (and to a lesser extent medical student training), and image interpretation. The body of work within image interpretation has been largely centered on perception and eye tracking. Although perception plays a large role in the process of radiology, the imaging study does not exist in a vacuum. Rather radiology is the synthesis of the visual perception, a solid knowledge in the related sciences, and the ability to link these to the clinical information. There has been little research done concerning how images are interpreted, and little that explores how novices or experts conceptualize what they are seeing within an image. This remains a large domain open for research.

The fact that medical education has been largely taught in silos, with little connection between subject content, it is not surprising that the vast majority of medical education research has been investigated in a similar manner. There have been minor attempts to correlate research across disciplines, however more work is necessary to build this integration and hopefully have a beneficial impact on medical education.

Establishing the Motivation of My Research

Despite the ongoing transformation in medical education curriculum, there still remains detachment in the correlation between imaging studies and physical examination findings. The newer medical school curriculum integration is attempting to bridge the gaps and foster relevant correlation of content, and maybe even research domains; however there are still elements of the curriculum that are taught in silos. It takes tremendous manpower and innovation to teach a fully integrated curriculum, and in my eighteen years of experience teaching medical students, the correlation of the physical examination and the patient's actual imaging studies (not just the written radiology report) seems to be lost and often treated as separate entities in patient diagnosis

Thus, I investigate the correlation of anatomy, physical examination and medical imaging and approach this from a cognitive framework. As a radiologist, I describe why this is important from an imaging perspective. The methodologies used in this study to achieve my goals are simulation, mixed methods and the clinical interview. Following a brief discussion of why this investigation impacts radiology, these methodologies will be explained and then describe how they were used in this investigation.

Radiology impact

Diagnostic medical imaging is one component of healthcare that has improved patient outcome in terms of increased life expectancy (Cutler & McClellan, 2001; Duszak, 2012; Lichtenburg, 2011). The knowledge of radiological anatomy is highly relevant to clinicians' daily work as imaging studies are becoming more central to clinical pathways and clinical decision-making (Pathiraja, Little, & Denison, 2014).

Not all physicians will need to understand and manage all details of imaging studies to the magnitude that radiologists do on a daily basis. However, there will be times when physicians, residents and medical students will need to interpret an imaging study or provide a preliminary read before a final radiology interpretation is made (Kiu, Bano, Barnes, & Khan, 2010; O'Brien, Cannarozzi, Torre, Mechaber, & Durning, 2008; Scheiner, Noto, & McCarten, 2002; Seagull, Bailey, Trout, Cohan & Lypson, 2014). Practicing physicians outside of radiology - family medicine, internal medicine, emergency medicine, and anesthesiologists - have recognized their deficiencies in interpreting medical imaging (O'Brien et al., 2008) and it has been reported that inaccurate initial assessments have impacted patient care in up to 11% of cases after an official staff radiologist assessment (Grosvenor, Verma, O'Brien, Entwisle, & Finlay, 2003; Hardy, Snaith, & Scally, 2013; O'Brien et al., 2008). When clinical issues remain unclear, there is a tendency to order imaging studies. To make the correlation of these imaging findings back to the clinical exam enhances the diagnostic ability of both components. When this correlation is not made, there is a potential lost diagnostic opportunity.

Simulation approach

Simulation embodies what it truly meant about learning experiences. Tyler (1949) defined a 'learning experience' as the 'interaction between learner and the external conditions in the environment to which he can react. Learning takes place through the active behavior of the student; it is what *he* does that he learns, not what the teacher does' (p. 63). Simulation allows a learner to be engaged with the learning experience (Pasquale, 2013). Teaching within medical simulation embodies critical thinking, problem-solving, and decision making skills, which are the goals of experiential learning (Pasquale, 2013).

In undergraduate medical education, simulation allows for opportunities to practice clinical conditions and skills that otherwise would not be possible in a safe protected environment. Simulation experiences address five important components lacking in today's medical education: (1) problems in clinical teaching, (2) the ever growing advancement in technologies for the diagnosis and management of diseases, (3) assessment of professional competence, (4) issues with medical errors, patient safety and team training, and (5) the role of deliberate practice (Issenberg, McGaghie, Petrusa, Gordon, & Scalese, 2005).

Because simulation can evoke these deeper levels of learning in an educational activity, it also provides an ideal framework to evaluate the cognitive processes that learners use to solve problems in a research setting. The ability to evaluate a learner within an active engagement setting allows for the ability to view their direct performance and also ask questions in the moment about their performance. Miller (1990) proposed four levels of assessment in medical education that could be actively evaluated in a simulation setting: (1) knows (knowledge) – has the ability to recall facts, principles and

theories; (2) knows how (competence) – ability to solve problems and describe procedures; (3) shows how (performance) – demonstrates skill in a controlled setting; and (4) does (action) – behavior in real practice.

Observing users while they view cross sectional images at a PACS (Picture Archiving and Communication System) workstation in radiology can lend insight about how users look at imaging studies and evaluate the anatomy on the images. Observing a learner while performing a physical exam on a patient will help to understand how a learner approaches specific anatomic areas related to a specific physical exam technique. However, to observe a user making a direct correlation of the physical exam and medical imaging would best be performed in a simulation setting that embodies both components. As part of this thesis, I developed a simulation tool that incorporates both the physical examination skills of touch and the visualization of cross sectional images to evaluate that direct correlation. Simulation and how it relates to my thesis will be discussed more in Chapter 2.

Mixed methods

Mixed methods research embodies both quantitative and qualitative data within a single study. Teddlie and Tashakkori (2011) describe mixed methods as methodological eclecticism. Where a study uses both qualitative and quantitative methods, because the best tool is what is chosen to answer a specific question. Mixed methods research allows for blending different knowledge claims, enquiry strategies, and methods (Maudsley, 2011).

Research in medicine is typically quantitative. It would have been very straightforward to capture the differences in novices and experts solely based on yes/no

answers, a probe placement location, or probe movements. However, this approach alone would not provide the reason why or the thoughts behind the participants actions. Combing these two approaches of quantitative physical data and the thoughts behind the actions (qualitative) assists to understand the similarities/differences in the novice and expert that one method of study could not attain alone. The choice of mixed methods allowed me to keep my research question central and use the particular method (qualitative or quantitative) best suited to that investigation.

Quantitative data was collected directly within the simulation tool, or direct answers to questions. However, the qualitative data that supported the quantitative data was collected through the development of a clinical interview. The use of mixed methods was implemented in my thesis will be further discussed in Chapters 3 and 4.

Development of the clinical interview

The foundation for developing a clinical interview is based on Piaget's early work with children and Vygotsky's work on zone of proximal development. The goals of the clinical interview are to depict spontaneous thought, identify thought processes, an establish context (Ginsberg, 1997).

The format of the interview gave important insight into the fundamental nature of thought (Ginsberg, 1981). Piaget developed a flexible method of questioning that allowed the exploration of children's thoughts and thus establishes cognitive competence (Ginsberg, 1981). Although Piaget originally developed this for children it has been adapted to gain insight into the thoughts of learners of all ages in educational settings and research.

Piaget's clinical interview had three aims: the discovery of cognitive activities, the identification of cognitive activities, and the evaluation of competence. The clinical interview is directed toward information–gathering (Posner and Getzog, 1982). Thereby the goal of the clinical interview is to obtain the nature and extent of an individual's knowledge about a particular domain (Posner and Getzog, 1982). Through the interview, relevant conceptions the individual maintains are identified and the relationship among them are explored.

The development of the clinical interview used in this thesis was an iterative process. Trials of the questions to ask the participants were piloted in two 4th year medical elective classes. The final version was used in this novice-expert study. A copy of the clinical interview and data worksheet are included in the appendix (1 and 2). The concept behind the interview was to explore different area of knowledge and thought process. The clinical interview developed contained four parts and provided a framework to begin the conversation. The first part was to prime the participants to begin thinking about how imaging and the physical examination relate. They were asked to recall a clinical case that involved in the interrelationship of the physical examination and any form of medical imaging. Now that they have had a chance to begin that line of thinking, the participants were introduced to the simulation tool. The next three parts of the clinical interview focused on using the simulation tool as they explored questions on knowledge construction, knowledge evaluation and scenario prediction as they related to the correlation of the physical exam and cross sectional imaging.

The format of the clinical interview helped maintain that the same core elements were collected from each participant. However, it did afford the flexibility to probe areas of

inquiry that arose during the dialogue. The core elements allowed for ease of comparison across participants. The use of the clinical interview and how it became integral to my information gathering about my participants will be discussed briefly in Chapter 2, and more fully in Chapter 3 and 4. The data developed from the clinical interview will be referred to as 'think aloud'.

The following chapters explore the three major components of my thesis project. The first section will discuss the development of the simulation tool that provides a direct correlation of the physical exam anatomy and medical imaging by the users. The next chapter will explore how this simulation tool was used to evaluate cognitive skills of novices and experts correlating the physical examination with medical imaging. Experts excelled in many of these skills. The last section, evaluates how novices and experts evaluate peer-based assessments within the simulation environment. Novices and experts have similar capacity to recognize correctness in peer-based assessments. This simulation tool can be used to evaluate the correlation of the physical examination and medical imaging within the novice-expert paradigm.

Chapter 2

Simulation can facilitate teaching the correlation of the physical exam and imaging anatomy

Abstract:

OBJECTIVE

Although both are important for making medical decisions and diagnoses, human anatomy associated with the physical exam and imaging studies are often taught separately and disconnectedly. Simulation can be used as a teaching method for correlating the anatomical components of the physical exam and imaging studies.

METHODS

A simulation tool was designed and built to support users in correlating imaging studies and the physical examination. Students were asked to use the simulation to translate between cross sectional imaging studies and the physical location of the presented anatomy.

RESULTS

The simulation tool provided opportunities for and built confidence in their abilities to correlate the physical examination and imaging studies. This study also identified strategies that the students use to make the correlation of physical examination and imaging studies.

CONCLUSION

Simulation can be used to correlate the knowledge base of anatomy and physical exam skills with imaging studies. It offers one method to teach this integration and provides some early insights into how students reason about anatomy in their introductory level of clinical medicine.

Introduction

Medical knowledge continues to exponentially expand. This expansion challenges educators and medical students to manage the required knowledge and skills necessary to function as competent clinicians (Johnson, Charchanti, & Troupis, 2012). The use and limitations of radiology imaging is one component of this knowledge expansion.

Although the clinical history and physical examination are the gateway to making the diagnosis, many diagnoses require additional medical studies, laboratory or imaging studies, to solidify a diagnosis. As such, the knowledge of radiological anatomy is highly relevant to clinicians' daily work as imaging studies are becoming more central to clinical pathways and clinical decision-making (Pathiraja, Little, & Denison, 2014).

Unfortunately, the current curriculum does not foster the correlation of these two disciplines, and historically this integration has not been valued. Most medical education curriculum focuses on teaching the physical exam and medical imaging separately, when in fact patient care can be improved by teaching them as an extension of the other. The skill of merging the physical exam and imaging studies is not easily attained, and often takes years to acquire and most medical schools do not offer opportunities to acquire these skills.

Even students notice the disconnection of the physical examination and medical imaging studies. During a discussion with a senior medical student he remarked that there have been opportunities on clinical rotations where correlation of the physical examination and imaging studies were made. However, he also expressed that this does not occur very often. He recounted a patient admitted for pneumonia after suffering a stroke. During that admission they used the patient's prior head CT and MRI to perform additional neurological testing that localized deficits on exam and then correlated those physical

findings with the medical images.

Additionally, when a student was asked how they make sense of integrating the physical examination and imaging studies, he first responded saying… 'umm, this is interesting' and after a minute or so of circling around non-connected thoughts, finally responded saying…. 'I can't say that I have put a ton of thought into it in the past'.

As this student's experience demonstrates, we need to figure out how to teach physical exam and imaging in a way that they are integrated and has value in patient care. The medical education community has yet to come up with good ways to do that. In this paper I will review how things are typically taught and suggest that simulation, as is used in other areas of medical education, could be a good place to turn for teaching this integration. Then I will describe a simulation tool I developed to accomplish this, and report on a study of its use with medical students to demonstrate its feasibility and utility for teaching this important integration of skills.

Literature Review

Most practicing doctors and medical school instructors would agree that integrating physical example with imaging is important. And yet, at this point in medical education, physical exam and medical imaging are taught independently, with little interest in helping students to develop these skills within the current course-based medical curriculum.

Teaching of Anatomy and the Physical Exam

The study of human anatomy has long been grounded in gross anatomy. It involves not only learning all the million anatomic structures in the body, but also the mastery of the

vocabulary and clinical reasoning skills (Savran, et al., 2015). Learning medical anatomy is a time intensive course, occupying over 100 hours of didactic content and laboratory experience in gross anatomical dissection (Albanese, 2010). Beyond the study of morphology or the structure and function of the body, anatomy is also the gateway into learning the geography of the biological entity (Guttmann, Drake, & Trelease, 2004).

This anatomical knowledge is taught through the experience of anatomical human dissection that involves the hands-on and personal or emotional aspects of the educational journey to become a physician (Granger, 2004). Some institutions have replaced or supplemented gross anatomy courses with innovative electronic teaching and learning tools (Jenkinson, 2009). These innovations have lead to great discussion about how anatomy should be taught. Some of these changes have been the result of curriculum reform and others due to financial reasons. Still others have suggested that although dissection is interactive, there is a need for more experiential learning in medical anatomy and propose a combination of various educational methods: dissection/prosection, interactive multimedia, procedural anatomy, surface and clinical anatomy (Asad & Nasir, 2015; Azer, 2013; Bergman et al., 2012; Dusseau, Knutson & Way, 2008), simulation (Hyde, Erolin & Ker, 2012) and medical imaging (Drake, 2007; Gunderman & Wilson, 2005; Jones, Olafson, & Sutin, 1978; Rizzolo, et al., 2010; Sugand, Abrahams, & Khurana, 2010; Wilson & Nava, 2010). All of these attempts have been to make the anatomy more relevant to the caring of patients.

Through coursework and study, medical students master the extensive content of human anatomy and some of the more common pathologies or conditions that are directly associated with anatomy or anatomical variants. However, the transition of this knowledge

to performing a physical examination is through skill-based courses. Most medical schools have introductory courses in the first year of medical school that teach the fundamentals of history taking and performing a physical examination. Most physical examination skills are learned through an apprenticeship practice (Verghese, Brady, Kapur, & Horowitz, 2011; Wenrich, et al., 2011). These early lessons prepare the student for the correct way to perform physical exam techniques. There is also exposure to normal and abnormal physical exams, thus enhancing the spectrum of variability to be expected in the normal physical examination.

Teaching of Imaging Skills

Alliance of Medical Student Educators in Radiology (AMSER) as an affinity group within the Association of University Radiologists (AUR) has developed a four-year radiology curriculum for medical students. The curriculum details core radiology topics, an organ based curriculum, diagnostic short lists and overall goals and objectives (Lewis & Shaffer, 2005; Lewis & Shaffer, 2010*)*. This content is specific for medical students, and should be based on their level of instruction for best integration.

Medical students should be exposed to radiology imaging as part of their training. This ranges from exposure to normal and abnormal imaging studies, different imaging modalities, risks of radiation, appropriate image ordering practices (Dillon & Slanetz, 2010) and the clinician-radiologist interaction as a consultant (Lewis & Shaffer, 2005; Lewis & Shaffer, 2010; Reddy, et al., 2015).

The methods and timing of medical student exposure to medical imaging content is variable from institution to institution (Afraq & McCall, 2002; Leschied, et al., 2013). There

is variable integration in the preclinical and clinical years, (Branstetter, Faix, Humphrey & Schumann, 2007; Collins, Dotti, & Albanese, 2002; Feigin, Magid, Smirniotopoulos & Carbognin, 2007) and variable instruction from radiologists and non-radiologists (Mirsadraee, Mankad, McCoubrie, Roberts & Kessle, 2012). Studies have demonstrated that radiologic anatomy averages about 5% of the total teaching time in medical anatomy courses (Ganske, Su, Loukas, & Shaffer, 2006). Imaging studies are often included in the didactic lectures as an extension of the anatomy and pathology course material (Leschied et al, 2013).

During clinical rotations, the medical students are exposed to common techniques to examine a chest x-ray, which is the most common imaging study that they will be exposed to in patient care (Jeffery, Goddard, Callaway, & Greenwood, 2003) and a required competency by the ACGME (Accreditation Council for Graduate Medical Education) (Eisen, Berger, Hegde & Schneider, 2006). These techniques are often taught on the internal medicine and surgical rotations (Eisen, et al., 2006). Most medical school programs have clinical rotations available for students in radiology, however often this is not a required rotation. Only 25-29% of medical schools require radiology as a clinical rotation (Leschied, et al., 2013; Eisen, et al., 2006).

Even on required rotations, students are often subjected to didactic lectures that increase their exposure to the types of imaging studies and their clinical uses, with limited access to active interpretation of imaging studies (Dmytriw, Mok, Gorelik, Kavanaugh, & Brown, 2015; Scheiner, Noto, & McCarten, 2002). The students often are observers of image study interpretation rather than an active participant. When students were surveyed about their radiology experiences in medical school, 63% believed it was

inadequate with 19% believing it is very inadequate (Dmytriw, et al., 2015). Despite this, the majority of medical students view radiology as an important component of healthcare.

Lack of integration

Based on the current teaching methods of anatomy and lack luster approach of most radiology education programs, despite the great opportunity for these two disciplines to interact, there is minimal integration. Historically, basic science anatomists have strictly taught anatomy in medical schools, with the vast majority of the curriculum taught through lectures and dissection (Ganske, Su, Loukas, Shaffer, 2006). In a survey of United States medical schools conducted by Ganske and colleagues (2006) they found that the course directors for medical anatomy were anatomists or cell biologists in 70% of schools, and radiologists compromise 1% of course directors. Eighty percent of schools use radiology images in their curriculum, but it comprises only 5% of the overall curriculum (Ganske, et al., 2006).

Radiology has a huge potential to have impact on anatomic basic science (Ganske, et al, 2006; Phillips, Smith & Straus, 2013). This limited role of imaging use in medical schools (5%) represents an undervaluation of the importance that radiology imaging can have on learning anatomy and providing life-long relevance to the anatomy.

Moreover, there have been an increase in radiology images being taught in U.S. medical schools over the years, however, there is less imaging taught by radiologists over time. Anatomists have become versed in looking at a set of imaging studies and often teach these findings to the medical students (Phillips, et al., 2013; Straus, et al., 2014). This provides a message that a radiologist's importance in interpreting imaging studies has

minimal value beyond what an anatomist can see. In doing so, there is a huge component missing about the clinical reasons how and why the imaging studies were obtained. The anatomists do not have the skill set to teach these facets of medicine and can not take the imaging beyond a strictly normal study to show derangements due to trauma, age, illness or neoplasia. Thus the integration of the radiologist in this capacity is much needed in the teaching of an integrated medical anatomy.

Often if medical students want to learn more about radiology they need to take an elective in their third or fourth years. Many schools do not have required radiology electives; only 10-25% in U.S. require radiology as an elective (Straus, et al., 2014). Thus, the integration of anatomy, medical physical exam and imaging is never taught or left to the students to figure out on their own.

Use of Simulation in Medical Education

There has been a resurgence of simulation in medical education with the curriculum reform that began in the 1990's. Three main factors have promoted the growth of simulation in medical education. One, more advanced simulation technologies have permitted devices with sufficient fidelity for realistic and meaningful engagement of learners. Two, reforms in medical education (Flexner report revisited (Cooke, Irby, & O'Brien, 2010) and the need to prepare undergraduates and beyond for the ever changing world of medical practice. Three, the widespread adoption of clinical governance to ensure we are providing the safest and highest level of care to patients. Thus simulation will impact risk management, lifelong learning, education and continuing education, quality

improvement and management of poor performance (Bradley, 2003; Bradley 2006; Friedman, 1995).

In medical education, it is difficult to ensure that every student is exposed to the same clinician experience, see the same spectrum of patients, and learn all the necessary skills (Lane, Slavin, & Ziv, 2001; Kalet, et al., 2012), thus simulation can offer students the breath of experiences and exposures during their training. Simulations can be simple to complex. They may involve simulated patient encounters (role playing or standardized patients), computer-based simulations (curriculum based skills like listening to heart sounds or reading EKGs)(Hyde, et al., 2012), to more complex high-technology simulations that evoke a complex clinical environment (cardiopulmonary resuscitation or invasive clinical procedures)(Lane, et al., 2001; Friedman, 1995; Murray, et al., 2002)

In undergraduate medical education, simulation allows for opportunities to practice clinical conditions and skills that otherwise would not be possible in a safe protected environment. High fidelity simulation addressed five important lacking components in medical education: (1) problems in clinical teaching, (2) the ever growing advancement in technologies for the diagnosis and management of diseases, (3) assessment of professional competence, (4) issues with medical errors, patient safety and team training, and (5) the role of deliberate practice (Issenberg, McGaghie, Petrusa, Gordon, & Scalese, 2005).

I propose that simulation can also be used in the beginning of medical education to integrate physical examination, medical anatomy and imaging studies, which have largely been taught in silos. This simulation tool, developed specifically to correlate medical imaging with the physical exam, provides the only direct correlation of these disciplines. Without this tool, this correlation is not possible in the current curriculum. With the

paradigm shift in medical education leaning towards competency-based assessment (Schumacher, Englander, & Carraccio, 2013), teaching methods that use and integrate imaging studies into patient care can be a pivotal assessment and also enhance the understanding of disease processes being presented in the curriculum.

Research Question:

Given the need in medical practice to integrate the physical examination and medical imaging, and the current disconnect between them in medical education, this work examines the question: Can simulation be used to integrate the anatomy present in the physical examination and medical imaging?

Design of Simulation Tool

To capitalize on the affordances of simulation and the ability to teach the integration of the physical exam with medical imaging, the first author designed and developed a simulation tool. Below we describe the design and use of the simulation. The simulation tool allows the user to control a calibrated CT scan over a human torso by moving a handheld probe over the torso and viewing the 'real-time' images on an iPad. The integration of the handheld probe controlling the images allows for direct correlation of the physical exam and the imaging at that location (figure 1).

Hardware design

The three-dimensional simulation instrument is a fiberglass human torso (armless torso from the head to the upper thigh) integrated with an iPAD interface that displays

medical human computed tomography (CT) images (figure 1). The CT image sets are navigated but moving a handheld probe over the body torso in the respective axial, sagittal or coronal planes (figure 2). The handheld probe is navigated through the Polhemus tracker connected to a laptop. Through a server connection interface, the laptop is connected to the iPad. This current interface requires a WiFi connection for the laptop and iPad to interact (figure 3).

The Polhemus tracker is a magnetic motion-tracking sensor that is able to measure position and orientation with six degrees of freedom (Polhemus, 2014). It allows for accurate measurements and the sensors are embeddable, thus not requiring a line of sight. An important quality for this simulation is that it is not affected by ambient light, although the medical images are best viewed with the overhead lights dimmed.

Images

The image set consists of an anonymized normal male of the abdomen and pelvis, which are displayed, in the axial, sagittal and coronal planes (figure 4). One image one of this contiguous set is initially shown to the participant on the integrated iPad to localize within the simulation torso. This anatomical region was chosen because it is accessible to the physical examination and also has important clinical and external landmarks that are commonly used to evaluate the underlying anatomy. CT imaging was used as the visual media because it portrays human anatomy with great detail and accuracy, thus allowing for greater clarity in correlation of imaging and physical examination. CT is also an imaging modality that most students have been exposed. Although some of the user interface in

this simulator resembles ultrasound imaging, most students have not had adequate ultrasound exposure and ultrasound imaging is a very user dependent modality.

Software

The CT images are presented on the iPad using the Unity gaming platform. Unity is an interactive 3D game development engine that supports multiple platforms and operating systems. The positioning and movements of the handheld probe are collected during the simulation events on the game interface on the iPad with ADAGE. ADAGE (Assessment Data Aggregator for Gaming Environments) is a software developed at the Game + Learning + Society Center at the University of Wisconsin - Madison with a partnership with The Learning Games Network and the Wisconsin Institutes of Discovery (Games Learning Society, 2012). ADAGE is a set of data collection and analysis tools on a shared, open source platform that uses big data techniques to transform clickstream data from game and design systems into formative feedback for usability testing and evidence of learning. This software allows for accessing in-game data that is a context-rich framework for using in learning analysis (Owen, Ramirez, Salmon & Halverson, 2014).

The ability to collect the localization in space allows for comparison to a targeted image. In addition, the ability to track probe movements in space by the user will allow access to not only to their thoughts (during think alouds) but also their actions in response to a question (localize the image within the body). Thee action movements not previously investigated, may allow insight into strategies or patterns of how users localize images.

Research Methods

Once the simulation was developed, built, and shown to be fully functional for users, I designed a study to test the feasibility and utility of the simulation device for coordinating anatomy and imaging in teaching medical students.

Participants

This IRB approved study was conducted with senior medical students in an advanced anatomy medical student elective at a large Midwestern medical school. The first group of students (n=11; 7 males, 4 females; average age 28 years with range 25-31) participated in a simulation exercise that included pre-test and post-test questions, and quantitative and qualitative simulation assessments on their ability to localize a CT image within the simulation device. A second similar group of senior medical students participated in the simulation exercise that included only quantitative and qualitative simulation data (n=17; 12 males, 5 females; average age 28 years with range 25-32).

Data collection

Pre- and post- surveys

I designed a survey to elicit the student's perception of how well they are able to interpret imaging studies, and also correlate those studies to physical exam findings. The surveys contained completion and 6-point Likert scale (bipolar scaling method) questions. A 6-point scale was used for the questions to force the responders to make a choice (Garland, 1991). The pre-survey included five items, which included three Likert questions. The post-survey included ten items, evenly divided between Likert questions

and open-ended questions. The open-ended questions asked in the pre- and post surveys, explored the student's comfort with looking at cross-sectional imaging and their confidence with relating the images to the physical exam and vice versa. Also, the survey data regarding the student satisfaction stands for their perceived utility of the simulation tool in the correlation of the physical exam and imaging studies

Simulation session data

After taking the pre-survey (if applicable), the researcher showed the student a normal CT image (called the test image) and asked him/her to use the probe to identify where on the simulation torso that image would be found. Once students placed the probe on the torso, they were allowed to move the probe until they felt the simulation image matched the image given by the researchers.

The simulation session data was used to identify how accurate students could correlate a CT image to its location in the physical body. The data from the simulations was recorded through the ADAGE computer program specifically developed for this simulation. The test image for each session was pre-assigned and thus has a coordinate axis in space. When the user placed the probe on the model and sets that location, that location is recorded as their beginning point. The computer then logs in the movements of the probe until the user completes the task. When the user is asked to process the location of an axial image in the body, the probe only records their superior and inferior movements. Likewise when the user is processing a sagittal image, the probe only records the left and right movements.

During the 30 to 45 minute simulation session, the participants were given an average of five (5) simulation events (range 3-8). Each participant was given a minimum of two axial images and one sagittal image (each image is referred to an event). The number of events each participant completed was adjusted based on student questions, ability and time.

Think Aloud data

While the participants were assessing the location of the image within the simulation body they were asked to verbalize their thoughts. Think aloud is a method of gathering information that provides rich verbal data about an individual's reasoning during a problem task. It provides the in-the-moment information about what is being concentrated on and how information is structured during a problem-solving task (Ericsson & Simon, 1980; Fonteyn, Kuipers & Grobe, 1993). This method provides valid data on the underlying thought processes that are occurring during the activity (Ericsson, 2006). Ericsson and Simon (1980) realized that when subjects are working under a heavy cognitive load, they tend to stop verbalizing or provide less complete verbalizations. In cases when participants neglect to verbalize, they are asked questions related to that event right after completing the task. The think alouds were audio-recorded for each participant and later transcribed.

Data analysis

Surveys

The survey questions asked on a 6-point Likert Scale were given a weighted value for each scale point. Because the Likert scale is ordinal, each Likert point was given a value (6= very comfortable/confident, to 1= not very comfortable/confident). The number of responders at the Likert point was multiplied by that value. All values for that question were summed, and then divided by the number of responders for the question. This allowed a comparison between different questions. These weighted averages were used to evaluate overall changes in their comfort and confidence in correlating imaging studies to the physical examination from the pre to post survey (Allen & Seaman, 2007; Hall, 2014). A chi-square analysis was performed on these differences.

Simulation Data

Each CT image presented to the user has a known coordinate location. The coordinate location of initial probe placement by the user, and their movements to get to the final target are recorded. The location of the probe with respect to the assigned image represents the best representation of where the user thought that image was located in the body. Secondly, the movement of the probe after initial placement provides insight into their understanding of the anatomy and its relationship to other structures. Chi-square analysis was performed on the initial probe placement for each simulation event to evaluate the localization success compared to the known location of the image.

Think aloud data

Qualitative methods are used to add richness to our understanding of what students are thinking as they navigate the simulation. Specifically, grounded theory is a method where data is explored for connections within and across individual participants that in turn are used to derive a theory from that data (Walker & Myrick, 2006). Qualitative methods allow for the exploration of the situation context and the interpretation of understanding in these interactions. Qualitative methods explore data that is not typically obtained with quantitative methods. Its strength's remain in the recognition of the significance of meaning, and the interpretation of this understanding to explain the phenomena (Maxwell, 2004). The fundamental understanding of the process relies on the context of the particular situation (Maxwell, 2004). It does not look for general patterns to explain a larger issue (cause–effect), but rather looks to explain what is occurring in the context of the situation (theory development).

Coding of these dialogue transcripts is one way that data can be broken down, compared, and then categorized. Coding is an iterative process that eventually leads to organization of the data into themes, essences, descriptions and theories (Walker & Myrick, 2006). The think aloud transcripts were iteratively analyzed for emerging codes that would support or refute the utility of using simulation as an educational tool for teaching physical examination correlation with imaging studies, and for student satisfaction with this tool.

Results

The results from this pilot test of two groups of students indicate the simulation can be used as an effective teaching method for correlating the physical examination and imaging studies. I describe the students' satisfaction with the simulation, increase in confidence in the practice, and the strategies they use to correlate anatomy and imaging.

Student satisfaction

When asked about their experience using the simulation, students had overwhelming positive comments about the simulation tool. During and after the simulation session the students made such positive statements like: 'I like the demonstration. I like that you can turn it on and show them exactly what they would see, where they are, but I guess also just walking through the different structures that you can see', 'it is nice just to -- yeah -- have some feedback as well from the machine at least', and 'It's interactive. … I thought that was really good'. The students really enjoyed to ability to interact with the imaging and a physical representation of a person. These interactions and comments indicate that simulation provides the students opportunity to make direct correlations with imaging and the body that they had not had in their prior medical education and experiences.

Not only did students like the simulation, they also suggest the simulation tool would be useful early in their academic trajectory. 'I think this would be helpful just even when you're starting to look at like imaging and CT scans and stuff, to be able to like place it on the body and figure out where you are. Because even like, I don't know, listening for -- like renal -- yeah bruits and stuff. Some people listen up here, some people listen

down here and like –'. The stimulation tool allows the students to confirm or adjust their clinical exam skills in a non-threating or high-stakes environment, which was not possible without this tool.

Increased confidence with correlating anatomy and medical images

The survey findings and responses from the students' also suggest the simulation may be effective for teaching the correlation of the physical examination and cross sectional images. On a 6-point Likert Scale the students were asked how comfortable or confident that they were looking at cross sectional imaging, relating the images back to the physical examination (PE) and relating physical findings back to imaging studies. Table 1 demonstrates the weighted average scores in the pre and post survey questions. The increase in average weighted score suggests that the students were more confident and comfortable with imaging and physical examination correlation after this simulation session.

All students had traditional course-based medical education. Prior to the simulation, the group had an average level (average score of 3 on 6-point scale, 6 is very confident and 1 is not confident at all; 4 students rating it a 4, 3 students rating it a 3 and 4 students rating it a 2) of comfort and confidence in relating imaging studies to the physical exam and vice versa. After this simulation session, their confidence and comfort increased to 5.09 (10 students rating it at 5, and 1 student rating it a 6). Analyzing these results demonstrated a significant improvement in the student comfort and confidence following the simulation (χ^2=8.25,p<0.01; χ^2=10.267, p< 0.01) respectively.

Student responses given during the open-ended survey questions provide insights as to how this simulation increased their confidence in correlation of imaging and physical examination. Here are several statements: 'this exercise helped me correlate the internal anatomy with the external surface and allowed me to appreciate relevant surface anatomical landmarks', 'It was helpful to actually think about the body sitting next to me while looking at the images', and "I was able to dynamically test my hypothesis regarding my ability to correlate images with the physical landmarks and confirm or refute my initial thoughts." Even in this single occurrence, the students found that this interactive experience allowed them to test their knowledge of correlating the images with important external landmarks and internal anatomy. This simulation tool allowed a unique opportunity for students to test their knowledge and skills that is not otherwise available.

Strategies students use to correlate anatomy and imaging

Palpation/Touching

Looking at the transcript data from the think aloud sessions, we can identify the kinds of activities students engaged in when correlating the physical exam with the images in the simulation. First, during the simulations students would often physically touch or palpate the model as if it were a human being (or patient). They would also locate physical examination or anatomic landmarks and relate those to the model and image. This student described how the simulation helped him, 'try to use your knowledge of surface anatomy to try and locate an image initially. Then I liked having the opportunity to adjust based on my knowledge of relative anatomy and what I was seeing'. This was a common strategy used by students to relate the imaging to the physical exam. They used the

physical exam findings (touching or landmarks) and the anatomy that they represented and used those references to localize the image that they were presented.

Scrolling

A second common strategy used by over half (6/11; 54.5% of students during the pretest survey) of the students was to be able to scroll through the images to locate anatomy. Scrolling through images was there preferred method to understand a cross sectional image set such as a CT scan. While performing one of the events, this student made this comment about how she prefers to look at cross sectional anatomy: 'so normally when I look at images and have questions on them, I scroll up an down to figure out where you are in the body. [Tone is a bit irritated, because she feels limited in not being able to move image to know where she is in the body] So normally I would scroll up and down, so looking at one slice makes it more difficult obviously.'

These verbalized comments may help to understand students' actions within the simulation data. Although the students were instructed to precisely localize the CT image presented to the physical location in the body, they tended to place the probe above or below the target (Table 2). The students expressed an underlying desire to want to scroll through images rather than locate one slice.

Students tend to significantly over-localize the anatomy presented depending on its location above or below the umbilicus (χ^2=20.41,p<0.0001). For example, when students are presented with an axial target image from the upper abdomen (liver, stomach, pancreas and spleen) above the umbilicus, the students will assess the location to be higher than presented in 76.5% (62/81 event trials). Similar but not statistically significant findings

were noted in the sagittal plane. Their initial localization of the anatomy may be close to the target. However, their reliance on a learned behavior of scrolling may limit the assessment of true direct correlation of the image and physical examination. Further investigation would need to be performed to identify the influence of scrolling has on their image localization in the body.

Other strategies

In addition, students would also use other sense making strategies to correlate the imaging and physical exam. These ranged from using a prior experience, relationship of adjacent anatomy, identifying the region of the body and describing a structure shape. These additional strategies and examples are outlined in Table 3.

These are the types of strategies that we might want students to develop skill with as they work on correlating physical exams and imaging studies. Further investigation is needed to determine which strategies are more successful in making the correlation of imaging and the physical exam, and which strategies we should encouragement fostering in our learners.

Discussion

The results presented above suggest that the simulation provides opportunity to correlate the physical exam with imaging studies that the students would not otherwise have in their traditional curriculum. But why might it work so well?

Medical imaging today provides a highly detailed picture of human anatomy. Medical decisions are made daily relying on the understanding of imaging anatomy and its

aberrations. The interpretation of medical images is largely a cognitive skill, however there are many components of this cognitive skill. These skills range from understanding the technical factors of the image, recognition of structures and anatomy, and problem solving involved in the particular clinical case presented (Manning, 2010). Interpreting medical images involves combining two skill sets. First is the ability to describe the objects presented on the imaging study and secondly is the ability to make sense of them with respect to the state of health or disease. This process takes years to master, but there are three ways simulation might helps.

One reason the simulation might support this mastery is that there appears to be a component of deliberate practice in the acquisition of expertise in correlating human anatomy to imaging anatomy, and vice versa. Ericsson (2004) argued that to acquire mastery of a new skill it was much more than an innate ability, but it required deliberate practice. In order for learner's engagement in deliberate practice to yield a desired educational outcome involves a controlled setting allowing for (1) repetitive performance of the intended skill in a focused setting, coupled with (2) rigorous assessment of skills, that provides (3) specific and informative feedback, thus (4) resulting in increasing skill performance (Ericsson, 2004; McGaghie, Issenberg, Petrusa & Scalese, 2006; Duviver, et al., 2011; Issenberg, et al., 2005).

The simulation offers all of these features. First, the simulation is set up that the user could repeatedly to test their skills on localizing an image within the physical torso. The user could practice the localization of images and relate back to the physical exam, or vice versa. Secondly, it provides rigorous skill assessment in that there is only one answer for each image. The location of the anatomy in this simulator is constant for each image set.

Thirdly, it provides specific and formative feedback in that each image is localized in space related to the torso. When the user matches the image to location a percentage score is provided to the user (something that was not stressed in this paper). Therefore it does provide correctness feedback to the user. Lastly, through repetitive practice the skill that is gain from this simulator could be transferred to other image sets or patient care (this has not been tested in this paper, but is a component for future evaluation).

Another reason the simulation might support student mastery is that it provides experiential learning which adult learners support as being a more effective method of learning. In experiential learning, learners build upon their knowledge through the interactions and experience with the learning environment (Pasquale, 2013). This supports the constructivists' view that the learner builds upon their prior cognitive structures (assimilation). Likewise, when the experience challenges existing cognitive structure, knowledge will be changed (accommodation) (Bradley, 2003). In this simulation session, students were not previously exposure to this particular situation. They drew upon their past knowledge of anatomy and clinical medicine to combine these principles. The simulation environment is intended to possess challenges to the learner's level of understanding, and still be perceived as safe. Often in simulation environments, the knowledge that the learner is exposed to draws upon prior knowledge but at a higher and more complex level, thus simulation also demonstrates spiraling as described by Bruner (Chauvin, 2012; Merriam, Caffarella, & Baumgartner, 2007).

Finally, the simulation may support mastery as learners take control of their learning and achievement objectives, and identifying resources or strategies that they believe will work best for them. This is one way to help facilitate scaffolding their

knowledge and developing life-long learners. The learners should be supported in their personal learning plans. The culmination of an effective simulation environment is to facilitate and engage learners in self-assessment in developing the skills in critical analysis, reflective practice, and self-responsibility for continuous learning and improvement.

Conclusion

This pilot study sought to investigate if a simulation device that incorporates medical imaging within a fabricated human torso can be used to teach the correlation of physical examination and medical imaging. The pilot was performed on senior medical students who should have a foundation level of understanding of human anatomy and medical imaging from their prior studies and clinical rotations. Through simulation events that involved locating the anatomy presented on a medical image within the physical body students were able to test their understanding of the physical exam and their knowledge of medical imaging anatomy. The students uniformly found that the experience enhanced their understanding of medical anatomy and expressed interest that this should be available to learners at all levels of medical education.

In addition to directly teaching this process to students, this simulation also provides us with the opportunity to begin to understand how students come to master this complex skill. At this point we know very little about people's understanding of anatomy and how this correlates in the understanding of medical imaging. Is it pure rote memorization, is it based on an understanding of the relationship to other adjacent structures or does it develop over years of practice in scaffolding with clinical situations?

This pilot study is just the beginning insight into how students make sense of medical imaging and its correlation with the physical exam. This tool, and the data we can collect about student thinking processes during learning, can help us better understand the incremental acquisition of anatomical knowledge and its application to clinical medicine. With this information, we can start to track the transition from expert to novice that must occur, as medical students become practicing clinicians, so that we can effectively and efficiently scaffold their learning in a system of ever increasing educational demands.

Figures

Figure 1: Components of simulation tool. A, iPad image viewer and control panel; B, handheld probe that moves image set in integration with the torso; C, torso model for physical integration with the image set and handheld probe; and D, mounted box housing all the controls.

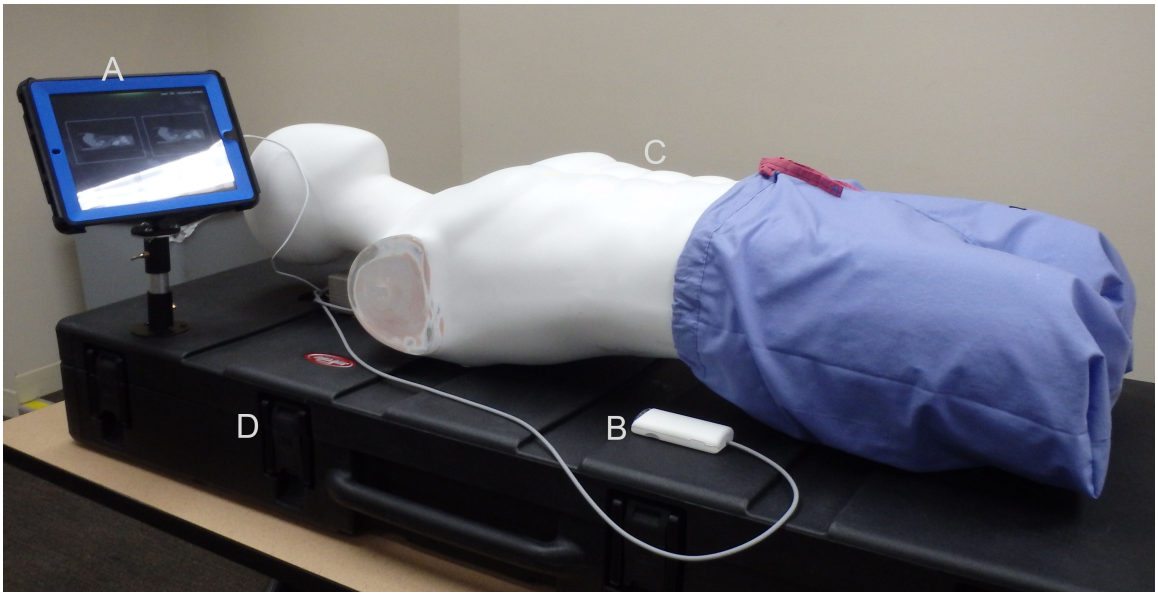


Figure 2: Simulation in use with handheld probe being used to direct movement of the images in relation to the location in the torso, and using touch controls on iPad image integration.

Figure 3: Schematic diagram of mechanics of the simulation tool. The laptop is connected to the images viewed on the iPad via WiFi connection. The probe connected to the Polhemus unit uses an electromagnetic field to provide the spatial reference of the probe and the cross sectional image on the iPad.

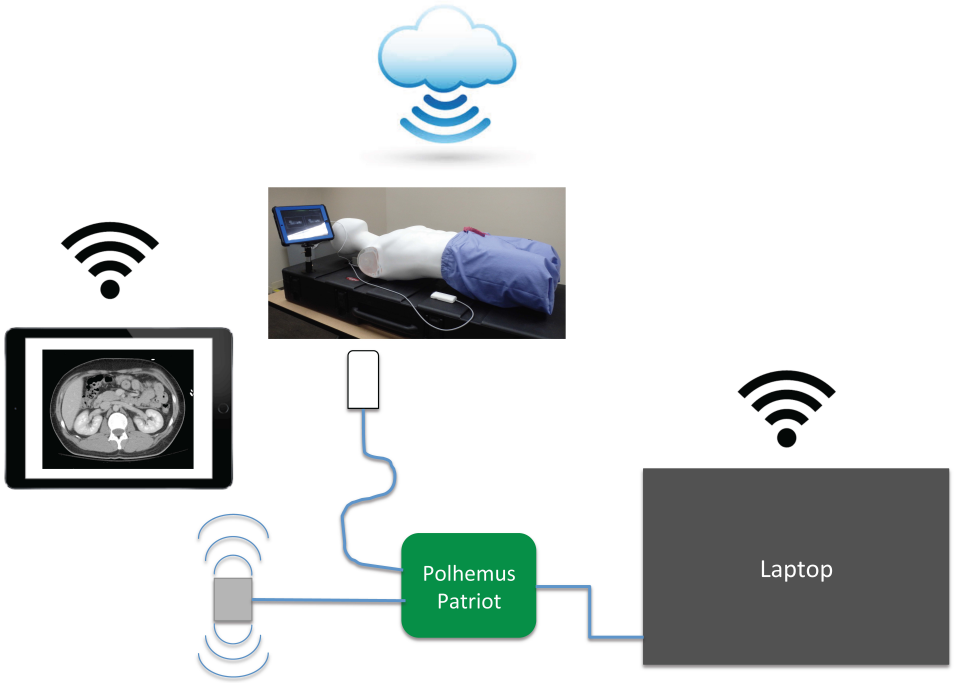
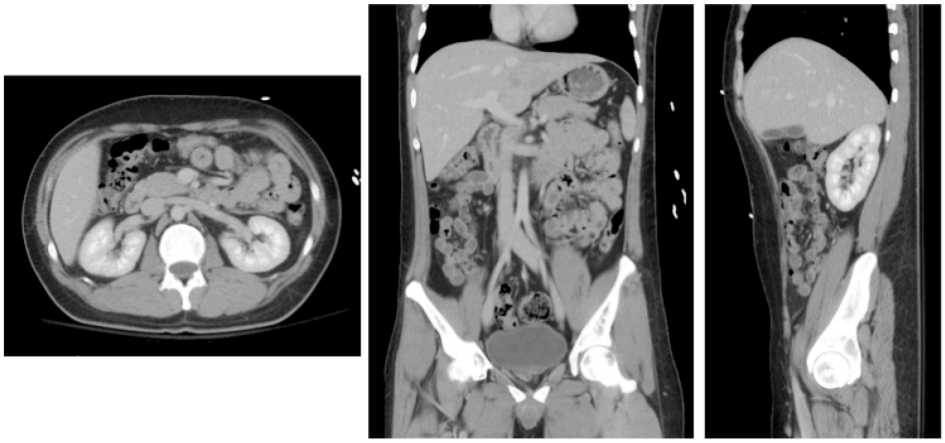


Figure 4: Example of cross sectional images of the abdomen that the students were presented during the simulation session.



Tables

Table 1: Average weighted survey scores of student's abilities to relate physical examination to imaging studies prior to and after the simulation exercise (results based on a 6-point Likert scale)

Survey	Content	Ave weighted score
Pre	Looking at cross sectional imaging	3.27
	Relating imaging to physical exam	3
	Relating physical exam to imaging	3
Post	Relating imaging to physical exam	5.09
	Will this simulation affect PE skills	4.64

Table 2: Localization positions of anatomy by the students relative to the location of the anatomy in the body and imaging plane.

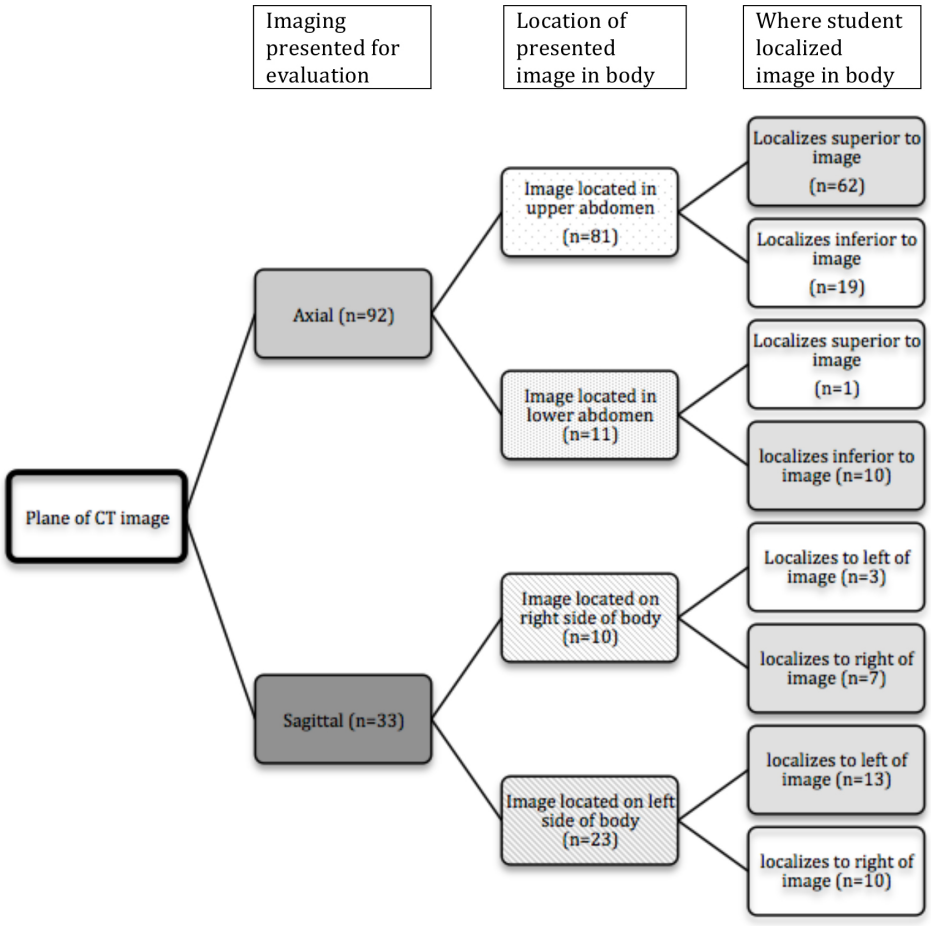


Table 3: Sense making strategies of students during simulation exercises

Sense making strategy	Student example
Spatial location/imaging plane	'this is an axial plane'
Identification of anatomy from prior experience	'Right away I look at the kidneys because I think I have a general idea in the body where they lie. Then I am glancing at the rest of the things to see if I find anything that really sticks out to me...like a landmark that I can use to kind of gauge where I am exactly. Um...'
Relationship of anatomy (higher or lower to image)	'So I guess I don't see kidneys so I know that I am not that low. I see the majority of the liver so I think I am higher in the abdomen. Seeing the view of the stomach with the pylorus…makes me think I am in the upper abdomen. I see the tip of the spleen so that makes me right where the ribs would end. So...those are kind of my thoughts of where I am on the body from head to toe.'
Region of the body (chest, abdomen or pelvis)	Note: the simulation only had one image set containing abdomen and pelvis, thus this may impact the amount of reference to this larger physical location reference.
Anatomical organs	The students would identify and name the organs that they could see within the imaging slice presented and then again when they were matching up the target image.
Type of imaging study	'Well I look at the type of scan, for me, because I am still learning' and ''What is the image? Is it an MRI or CT? So, kind of fuzzy, not as clear as an MRI, so it is probably a CT.'
Describe structure shape	'then once I was coming down to match the image it was more...does the curves look the same in the liver, was the air...dark spots..'
Discuss color differences in image (white/black/grey)	'so this looks like….there is this black piece here…..there are two black spots there' and 'I like looking at the really hypodense areas, the dark areas, and hyperintense areas, really bright.'

Chapter 3

Expert and novice differences in correlating anatomy and imaging

Abstract

PURPOSE

The ability to correlate anatomical knowledge and medical imaging is foundational to radiology. Experts do this well, but we have little understanding of how they learn to do so. Even more problematic, we don't know how novices transition into being experts. Knowing where our novice's understanding begins in making this correlation can aid us in helping them transition to being experts.

MATERIALS AND METHODS

Ten radiologist experts (average age 47.4 years with 13.5 average years of experience; 9 males, 1 female) and 11 senior medical student novices (average age 27.4 years; 5 males, 6 females) performed a simulation localizing axial and sagittal computed tomography images within a human simulation torso. This study was IRB approved. Quantitative simulation data was collected on image orientation, time, and correctness of localization. The simulation data was assessed with one-way and two-way ANOVAs. Additionally, participants engaged in qualitative think-alouds about their reasoning during the simulation sessions. The transcripts were coded and assessed for emerging themes. Chi-square analysis was performed on the qualitative codes. Significance was assessed at $p<0.05$.

RESULTS

Consistent with prior literature on expert-novice differences, experts are significantly faster at making decisions on medical imaging than novices ($p<0.001$). When localizing an image in the body, experts rely on organ substructures ($p<0.0001$) whereas novices rely heavily on size or amount of an organ in the image ($p<0.001$). Experts are more likely to use the correct terminology ($p<0.001$), whereas novices are more likely to misinterpret the anatomy ($p=0.002$) and use non-anatomic descriptive cues (color, blobs, patterns) to describe what they are viewing ($p=0.004$). Experts notice patterns on medical imaging not common to novices. When performing fine-tuning adjustments during localization, experts isolate a structure with a narrow zone of change ($p<0.001$), compared to novices who use the shape or size of an organ ($p<0.001$) or trial and error methods ($p<0.001$) when performing the same tasks.

CONCLUSION

There are expert-novice differences in image processing and correlation with anatomy. Specifically, the cognitive processing of experts and novices is different with respect to meaningful patterns, organized content knowledge and the flexibility of retrieval. The study of novice-expert provides meaningful understanding of their knowledge differences and insights to develop educational programs that promote novice-expert advancement.

Introduction

The knowledge of radiological anatomy is highly relevant to clinicians' daily work as imaging studies are becoming more central to clinical pathways and clinical decision-making (Pathiraja, Little, & Denison, 2014). With the growing use of imaging studies (Cutler & McClellan, 2001; Duszak, 2012), being able to correlate the elements of the physical exam and the medical imaging studies is an important clinical skill for physicians to have in the care of patients.

For example, as a radiologist, I use the correlation of imaging anatomy and physical anatomy on a daily basis. Breast cancer patients will often have image-guided biopsies of both a breast lesion and axillary node prior to getting neoadjuvant chemotherapy. Many times the tumor lesions are no longer visible after the chemotherapy, but the markers placed at the time of image-guided biopsy remain visible. In the breast, the visible biopsy markers can be localized with mammography. However, the markers placed in the axially nodes may not be visible to be localized by ultrasound or mammography. Prior cross-sectional imaging studies of the patient's (CT or MR) are then used to guide the surgeon to the location of the biopsy markers. This involves using prior imaging studies, the surrounding anatomical structures and how they relate to the external skin surface. This is challenging and complex but is a common practice for patient's whose disease requires coordinated care of the radiologist and surgeon to aid in removal of the entire tumor.

As the above example demonstrates, interpreting radiology images and coordinating them with anatomy is a complex skill. It involves combining the information from visual pattern recognition, anatomical knowledge, knowledge of pathological processes, and patient-specific information (Rubin, 1989). It goes beyond the task of

searching, interpreting and reading images and is the combination of perceptual and cognitive skills (Nodine & Krupinski, 1998).

How can we teach medical students this complexity so that they can perform it in their practice? Rubin (1989) postulated that the cognitive sciences might help to provide insights into how radiologists interpret imaging studies and thus provide resources to trainees in radiology. To effectively teach students, cognitive science suggests we should know the extremes of this skill – the final state of this skill (what expertise looks like) and where students start (the novice state) (Bransford, Brown, & Cocking, 2000; Chi, 2006a, 2011; Dreyfus & Dreysfus, 1980; Lesgold, 1983). Understanding how experts reason through problems can provide a model toward which an effective educational process can aim. In addition, understanding how novices engage in similar situations provides information about the skills and knowledge upon which instruction can be based.

Unfortunately, at this point we know very little how either experts or novices coordinate radiology and anatomy. Much research has focused on the perception sciences of radiology (Faruque, Rubin, Beaulieu, & Napel, 2012; Kok, De Bruin, & Robben, & Van Merrienboer, 2012; Krupinski, 1996, 2000a; Krupinski, Berger, Dallas, & Roehrig, 2003; Kumazawa, et al., 2008; Kundel, 2006; Kundel & La Follette, 1972; Kundel, Nodine, & Krupinski, 1989; Mehmood, et al., 2013; Mello-Thomas, 2006). However that work still has not examined the cognitive processes required to coordinate this with the physical exam. In this work I will begin to address this gap by exploring the thought processes of how novices and experts make correlations between anatomy and medical imaging using a specially designed simulation tool.

I begin by reviewing existing literature from cognitive science comparing various attributes of novices and experts across domains. I describe a study I designed to explore expert-novice differences in the domain of radiology, specifically the correlation of medical images to the physical body within a simulation tool. Then I present the findings of how novices and experts differ in these tasks, organizing the findings around the features of expertise identified in the broader literature. Finally, I conclude with describing implications for how we can teach this correlation of the physical exam and imaging studies to our learners.

Literature Review

We will begin this literature review with a discussion about constructivism. Constructivism emphasizes acquisition through knowledge construction (building upon prior knowledge). The learner is building and transforming the knowledge (Applefield, Huber, & Moallem, 2000). These are similar to the underpinnings in novice-expert research paradigm. Thus it is important to present the background of constructivism and later build upon these concepts by exploring the differences of novices and experts.

Constructivism

Constructivism views learning as a person's building upon knowledge in a way that makes sense to them. The meaning of this knowledge is dependent upon previous experience and the current exposure to knowledge (Bodner, 1986; Merriam, Caffarella, & Baumgartner, 2007; Schunk, 2012). Piaget as a pioneer of the constructivist viewpoint of

learning (Piaget, 1929) viewed this construction of prior knowledge through assimilation and accommodation (Piaget, 1964).

Drawing upon constructivism, all humans come to an educational experience with prior knowledge. This prior knowledge influences what they notice about their environment, how they organize this new information and how they interpret it. These factors will influence their ability to remember, reason, solve problems, and acquire new knowledge (Bransford, et al., 2000).

According to constructivism, people's prior knowledge is what is ultimately transformed into correct knowledge and understanding (Daley, 2000). People are not blank slates. It is important in any educational experience to know the existing knowledge of the learners. Building upon prior knowledge is the framework of the novice-expert paradigm, which is grounded in constructivism.

Novice-Expert Differences in Cognitive Science

With respect to the cognitive sciences, there are several areas that we will focus our discussion. The expert-novice paradigm is an area of growing research, it is important o understand the beginning and end points of a learning objective, and expert and novices are domain specific and different from one another.

The viewpoint that learners do not begin as empty slates has lead to a body of research comparing experts and novices in different domains. In the twentieth century, there was increasing interest by cognitive psychologists concerning the internal mechanisms responsible for attaining expert performance. Ericsson (1996) investigated

expert performance, which he described as 'consistent, measurable, and reproducible performance' (p. 3) by top performers in their domain.

There are two main ways researchers study expertise (Chi, 2006b). The first is to study expertise within their domain and to understand how they perform. The goal being to understand their superior performance, so the methods of assessment need to accurately portray their expertise. This method looks at how different these few are from the multitudes, and makes the assumption that there is something inherently unique about these individuals (absolute approach). The second is to approach the study of expertise by comparing experts to novices. This approach makes the assumption that expertise is an attainable goal by novices. Chi (2006b, 2011) calls this later approach the relative approach. An expert in this context is someone that is more advanced, and measured within the domain of his or her expertise. The goal of relative approach is to understand how experts excel, such that others can facilitate and/or accelerate to a level of expertise.

This leads to the concept of a continuum from novice to expert. A continuum opens up the assumption that expertise can be attained. Therefore the goal is to understand how less skilled or experienced (novices) person can become more skilled. Novices can build upon their prior knowledge to become experts. Thus it is important to know what knowledge our learners have at the beginning. Assessing only final knowledge, such as a test at the end of a unit or course, does not measure the knowledge gained, but only measures the content of knowledge on that particular instrument/evaluation.

Expertise is domain specific. An expert is someone that excels within a domain. Experts possess an organized body of knowledge that can be effortlessly accessed and used (Glaser & Chi, 1988). In contrast, a novice is someone that is completely new to an area and

who does not possess a significant amount of pre-existing knowledge of that discipline (Dunphy & Williamson, 2004). These two groups can be viewed as extremes of mastery.

Expert-novice differences in medical education research

What are the differences between experts and novices in the field of medicine? Medical education research has looked at expert-novice differences and the research can be divided into two domains: cognitive knowledge (amount of knowledge) and reasoning (how they approach problems) (Norman, Eva, Brooks, & Hamstra, 2006; Schmidt and Boshuizen, 1993). Experts and novices differ in each of these domains.

Cognitive knowledge

Researchers identified three broad areas of cognitive knowledge that define medical expertise: casual knowledge, analytical knowledge, and experiential knowledge. Casual knowledge is the understanding of the basic mechanisms of science. Schmidt and Boshuizen (1993) examined this knowledge by testing recall of clinical case information from novices, intermediates, and experts. The intermediates recalled more information than novices or experts in early experiments and later experiments (van de Wiel, Schmidt, & Boshuizen, 1998) demonstrated a linear increase in recall with expertise. Overall, the experts processed the clinical information using knowledge structures that were distinctly different from the novices and intermediates, suggesting that expertise is more than just recall of content (casual knowledge). Furthermore, the role of basic science knowledge plays little role in expertise (Norman, et al., 2006), but rather this knowledge is chunked

(Schmidt & Rikers, 2007; Rikers, Loyens & Schmidt, 2004; deBruin, Schmidt, & Rikers, 2005), such that it can be called upon when approaching a difficult problem.

Analytical knowledge involves use of signs and symptoms to make a diagnosis. Students can be taught the signs and symptom basis of disease (semantic qualifiers), however it did not have a great impact on their impact of diagnostic accuracy (Nendaz & Bordage, 2002; Crowley, Naus, Stewart, & Friedman, 2003). Expertise then is more than knowing the signs and symptoms of disease; students must put these together with prior knowledge. Bordage (2007) demonstrated that it was not just the use of these sematic qualifiers (signs and symptoms), but also the number of them (twice as many by experts) that lead to expertise and being a better diagnostician. Experts use the semantic qualifiers more frequently, and have a more diversified set (Bordage, 2007; Bowen, 2006; Crowley, et al., 2003).

Lastly, Norman and colleagues (2006) define experiential knowledge as the accumulation of prior cases that comes with experience. Experts have accumulated each prior experience as a category, and this learned experience is retrievable and provides support for the categorization of new similar cases. This process is not analytical or conscious, and thus may be stimulated by similarity based on features that are objectively irrelevant to the category. Accumulating this experiential knowledge may be part of the 10,000 hours of deliberate practice needed to become an expert as discussed by Ericsson (2004) (Nodine & Mello-Thoms, 2010). Novices limited exposure to experiential learning within a given domain imposes limitations on the development of cognitive knowledge. Kalet and colleagues (2012) investigated different types of interactivity in an abdominal exam module and evaluated performance of an abdominal examination on a standardized

patient. They found that enhanced engagement with the material, versus the students who observed, was associated with improvements in performing the clinical examination. Thus learning through active engagement assists novices cognitively in the construction of knowledge.

Reasoning

In contrast to cognitive knowledge, reasoning in medical education is the ability to draw information from the patient, medical records, laboratory and imaging studies, and other health care providers and combine that with cognitive knowledge (casual, analytical and experiential) to conclude upon a diagnosis (Norman, 2005; Schmidt & Boshuizen, 1993; Crowley, Naus, & Friedman, 2001). There have been several postulated ways that this occurs from hypothetico-deductive, pattern recognition, forward reasoning, and semantic qualifiers to name a few. No single reasoning process defines expertise; rather expertise in clinical reasoning appears to be a consequence of an extensive and multidimensional knowledge base (Norman, 2005).

Expert-Novice Differences in Radiology

Most of expertise in medicine has been focused on expertise in internal medicine. Radiology differs from other areas of medicine in that it is highly reliant on visual input. It involves a substantial perceptual component, formalized medical knowledge and knowledge gained in clinical experience. It also involves integrating bodies of medical knowledge that have distinct structures to them, including anatomy, physiology, pathology and the projective geometry of radiography (medical physics) (Lesgold, et al., 1988).

However, radiology also differs as a discipline of expertise as it needs to answer three questions with every image interpretation: what is it, where is it, and what is it that makes it what it is (Grill-Spector & Kanwisher, 2005).

There have been some studies examining expertise in radiology, but many of them have focused solely on perceptual skills. For example, Berbaum and colleagues (1990) tested the perceptual phenomena of 'satisfaction of search' by adding simulated nodules distractors on chest x-rays. Similarly, eye-tracking patterns were studied and demonstrated that it varies with level of expertise (Mello-Thoms, 2003; Mello-Thoms, et al., 2002; Voisin, Yoon, Tourassi, Morin-Ducote, & Hudson, 2013). Analogously, different search patterns with respect to visual attention and correction interpretation have been seen in experts and novices (Krupinski, 2000b; Mello-Thoms, 2003; Mello-Thoms & Hardesty, 2005). These studies have looked only at the perception piece of radiology – not its full complexity.

Far fewer studies have examined how expert radiologists analyze imaging studies and relate to clinical situation. One study maintains that expertise is an accumulation of specialized schemata that are sensitive to specific disease states (where a schemata is the perception of a finding that is dependent upon the other surrounding indirect evidence)(Lesgold, Feltovich, Glaser, & Wang, 1981; Lesgold, 1983). The other study asserts that expertise is the development of feature lists with accurate status values and appropriate combining weights to make a diagnosis (Getty, Pickett, D'Orsi, & Swets, 1988).

These studies examining expertise in coordinating imaging and the physical exam are important because of the specialized nature of radiological expertise. However, research from cognitive science that looks at expertise across domains also indicates there

are generic features of expertise that are likely present in some form in radiological expertise. This study looked at the ways in which the general features of expertise show up in radiological expertise.

General Features of Expertise

Synthesizing studies over the last several decades, Bransford and colleagues (2000) identified a number of principles of an expert's knowledge that may have implications for learning. These principles are assumed to apply across domains of expertise.

1. Experts have a great deal of ***content knowledge structured around organization***. The knowledge of experts is organized around core concepts or big ideas, and is not a list of facts or formulas within a domain (Bransford, Brown & Cocking, 2000; Glaser, 1984). Experts organize knowledge around large topics, and their ability to store concepts as 'chunks' of memory has the appearance of 'knowing more'. These chunks of memory make it more efficient to draw upon information and have more information to relate to the specific problem (Chi, Feltovich, & Glaser, 1981).

For example, Chi, Feltovich and Glaser (1981) asked expert and novices to sort 24 physics problems into groups based on similarities of solution. The expert's categorized the problems into types defined by major physics principles that will be used in the solution. Novices, on the other hand, categorized them into types defined by the items contained in the problem statement.

Within medical education, Grosswald (1992) studied expert and novices in obstetrics and gynecology by having participants solve three medical cases. The experts combined the copious information they gathered from a systematic approach to solving the

problem into meaningful chunks leading to the diagnosis. In contrast, novices tended to focus on only a single diagnosis, and actions were made to directly confirm that diagnosis in a minimal amount of steps. They looked at only information that would support the initial diagnosis, and disregarded the rest of the information, perceiving pieces rather than groups of information. The experts' organization of knowledge led to a more accurate and complete diagnosis.

*2. Experts **recognize complex patterns** that are not noticed by novices.* Experts are more likely to recognize meaningful configurations in data and realize the implications of these situations (Bransford, Brown & Cocking, 2000). Expert's ability to recognize meaningful patterns stems from their ability to chunk information into familiar patterns. Novices do not use this chunking strategy as they lack an ordered, highly organized structure for the domain (Miller, 1956).

In deGroot's work with master chess players, he noted that master chess players were able to recognize meaningful chess configurations and strategies, thus allowing them to make superior moves when executing a game (Bransford, et al., 2000). In the domain of medical education, Lesgold and colleagues (1981) presented novices and experts a chest x-ray with atelectasis and asked them to interpret. The experts were able to make the diagnosis of atelectasis, but also noted patterns of disease on the chest x-ray (emphysema and chronic lung disease) that limited the diagnosis of the otherwise ambiguous imaging findings. In both of these examples, experts were able to perceive complex patterns that for the novices did not come to awareness.

*3. Expert knowledge exhibits **situational applicability**.* Beyond an experts' acquired knowledge, they have the ability to retrieve the knowledge appropriate for the task. That

is, expert knowledge is 'conditionalized' – the knowledge is dependent on a set of circumstances. Conditionalized knowledge enables a person to access and apply domain-specific knowledge to a specific problem (VanSickle & Hoge, 1991). It is critical to know what problems/situations require what knowledge. Experts can retrieve knowledge that is relevant to the particular task (Bransford, et al., 2000; VanSickle & Hoge, 1991).

Master chess players when looking at a chessboard and reading it to make a move, will only consider a subset of possible moves. However that subset of moves is superior compared to novice players (Bransford, et al., 2000). The master chess players retrieve only the moves that are relevant to that situation.

For example, in medical education Lesgold and colleagues (1988) presented experts and novices a chest x-ray, without clinical history; of a patient whom long ago had a portion of the right upper lung removed. The expert quickly noticed that there were signs of chronic collapsed lung, which shifted chest organs to fill the void. Experts knew to apply knowledge of prior chest x-rays with surgical intervention and not fall to the more novice assessment of atelectasis causing the alteration seen on the chest x-ray.

4. The organization of expert knowledge reflects deep understanding and it can be ***flexibly retrieved with little effort.*** They are able to effortlessly retrieve knowledge to apply to a problem. This automaticity frees up the ability to attend to portions of the problem that need more attention. Not all problems will be solved faster (but this does afford the possibility it can be), however the fluency is important for understanding and recognizing problems in a domain (Bransford, et al., 2000). This fluency places fewer demands on conscious attention, which has limited capacity to solve a problem.

This automaticity of experts can be critical in making fast and accurate responses to critical stimuli (Gonzalez & Thomas, 2008). Military pilots while flying, without conscious awareness, need to distinguish the characteristics of friendly aircraft from the enemy and make the appropriate response in an instant (Gonzalez & Thomas, 2008). In Lesgold and colleagues (1988) atelectasis case, experts had more refined schemata allowing them to make finer discriminations to make the diagnosis of atelectasis over the novice conclusions of congestive heart failure or tumor. Their refined schemata allowed them to make subtle distinctions that were important in making the correct conclusion.

5. Lastly, experts ***easily adapt to a new situation****.* They can consider many more possibilities and vary their approach compared to the less experienced novice. They can 'act in the moment'. They approach new problems to use their expertise and look at it as an opportunity to expand their current level of expertise. They continually try to move beyond what they know and question their current level of expertise (Bransford, et al., 2000; Chi, 2011; Ericsson, 2007). They strive to do things better, not just more efficient. For example, a London cab driven, as an expert at knowing all the roads and routes of travel in London, needs to take his client to a destination. A road is newly under construction; the cab driver must quickly call upon other solutions to get his client to their final destination in a timely manner.

A qualitative study by Mahant and colleagues (2012) that explored the nature of excellence in top-rated academic clinicians identified three themes of excellence: core philosophy, deliberate activities, and everyday practice. Within everyday practice, these clinicians were noted to adapt the necessary skills to unique situations. They were able to adapt to and accommodate every different kind of patient and clinical problem. This was

described as the ability to recognize ones own limits and draw on expertise of others, but also to be able to solve problems and think creativity in new situations.

Research Question:

The history of expert-novice studies in medical education and radiology coupled with Bransford and his colleagues' (2000) generic features of expertise provide the starting point for my work. Specifically, in this expert-novice study, I seek to understand what features of expertise - as identified by Bransford and colleagues - look like in a specific sub-domain of radiology. I explore the specific context of correlating axial and sagittal CT images with the anatomy of the physical examination by asking:

What differences exist between experts and novices in correlating anatomy and cross sectional imaging?

Methods

To investigate this research question, expert and novice volunteers participated in correlating cross sectional medical imaging to the physical exam within the context of a simulation tool. Qualitative and quantitative data was collected during each simulation session with prescribed tasks. The data provides insight into the features of expertise in this sub-domain of radiology.

Simulation Tool

Since it would be difficult to collect controlled data of radiologists and novices looking at actual images in a clinical setting, we used a simulation tool that was developed for these types of scenario evaluations. I developed a simulation tool to examine expert-novice performance in this context. The simulation tool allows the user to control a calibrated CT scan of a human torso by moving a handheld probe over the torso and viewing the 'real-time' images on an iPad. The integration of the handheld probe controlling the images allows for direct correlation of the physical exam and the imaging at that location (figure 1 and 2). The image set consists of an anonymized normal male of the abdomen and pelvis. This anatomical region was chosen because it is accessible to the physical examination, and has important clinical and external landmarks that are commonly used to evaluate the underlying anatomy.

Participants

Ten radiologist experts who interpret abdominal imaging as part of their practice (average age 47.4 years with 13.5 average years of experience; 9 males, 1 female) were invited to participate. In this study, experts were defined because of their skill and years of training in evaluating abdominal cross sectional imaging. The novices were defined by their degree of back ground knowledge, as someone that has some exposure to medicine and medical imaging, but they have limited knowledge, senior medical students.

Given that research shows that expertise is heavily domain-dependent, radiologists who interpret abdominal imaging were chosen to exclude non-expert bias from radiologists that were sub-specialized in other fields such as neuroradiology,

musculoskeletal, chest or breast imaging. Six of the radiologists were community based and four radiologists were academic based within the same large Midwest University radiology practice group.

Eleven senior medical students from a large Midwest medical school (average age 27.4 years; 5 males, 6 females) were invited using a snowball invitation method (Atkinson & Flint, 2001; Biernacki & Waldorf, 1981; Cohen, Manion, & Morrison, 2007). Students from the author's advanced anatomy course who participated in a pilot study using the simulation tool invited peers to participate. These potential participants' contact information were sent to an administrative assistant who arranged the session times and dates. These students represent the novice group for the study. The students were fourth year medical students.

These novices were enrolled in a traditional course-based curriculum with a strong reliance on multiple-choice tests. The medical school curriculum was divided into two years of basic science and two years to clinical based medical rotations. They were chosen as the novice group because they have enough foundational knowledge of the basic sciences (including anatomy) and the applied clinical experiences to engage with the study tasks, but not enough knowledge and skill yet to be considered experts.

This study was IRB approved. All participants, experts and novices, were compensated with a university logo travel mug.

Study Activities

At the start of the study, participants were each given the same sequence of CT images (figure 3 and 4) to locate in the body. The CT images consisted of a practice case

(not used in data analysis), then two axial images, a sagittal image, two more axial images and lastly one more sagittal image. The same image set was used for all participants to allow for comparisons across levels of expertise.

During each of the events, the participant was presented with a CT image. They were given instructions to use the hand-held probe to localize that slice to the best of their ability within the human simulation torso. Once localized, they logged in that position of the probe by tapping a button on the iPad. At that point they able to see their success in localization with side by side images on the iPad screen of the target image and where they localized the image to be in the body. If they are not exact, they were instructed to match the two images by moving the probe in the appropriate direction. They finalized this step by tapping a button on the iPad.

Data collected

Simulation session data

The simulation session data was used to identify how accurately the participants correlate a CT image to its location in the physical body. The data from the simulations was recorded through the ADAGE computer program specifically developed for this simulation. The images for each session were pre-assigned and thus have a coordinate axis in space. When the user places the probe on the model and sets that location, that location is recorded as their beginning point. The computer then logs in the movements of the probe until the user completes the task. When the user is asked to process the location of an axial image in the body, the probe only records their superior and inferior movements. Likewise

when the user is processing a sagittal image, the probe only records the left and right movements.

In addition to location, the simulation session recorded continuous time points. However, only three time points were used for analysis in this study: the moment the participant saw the target image, the moment when the participant localized target image with the probe on the simulation torso, and lastly the moment when they completed the final matching of their original probe location to the target image.

Think Aloud data

While the participants were locating the position of the image within the simulation body they were asked to verbalize their thoughts. Think aloud is a method of gathering information that provides rich verbal data about an individual's reasoning during a problem task. It provides the in-the-moment information about what is being concentrated on and how information is structured during a problem-solving task (Fonteyn, Kuipers & Grobe, 1993; Ericsson & Simon, 1980). This method provides valid data on the underlying thought processes that are occurring during the activity (Ericsson, 2006).

Although think aloud can provide access to cognition, Ericsson and Simon (1980) found that when subjects are working under a heavy cognitive load, they tend to stop verbalizing or provide less complete verbalizations. To ensure capturing complete data, in situations when participants neglect to verbalize, they are asked questions related to that event right after completing the task. The think alouds were audio-recorded for each participant and later transcribed.

Data Analysis

Quantitative Data Analysis

The simulation data was assessed with one-way and two-way ANOVAs. One-way ANOVAs were used to assess for mean performance differences of experts and novices within different time segments of the simulation (preset to ready, ready to completion, preset to visual threshold, and ready to visual threshold) (Table 1). Here, 'preset' is the time point when the participant is presented with the target image on the iPad, 'ready' is when they have placed the probe on the simulation torso where they think that target image is located, and 'completion' is when they have their final adjustments to match up the target image. To allow for any internal lag in the probe movements and the visualized iPad image, a 'visual threshold' was developed. This is the spatial range of where the probe could be on the simulation torso and have the represented target on the iPad screen. This visual threshold was developed for all tested targets by the first author. Each target was tested twice on the simulation tool and the average threshold range was calculated.

Two-way ANOVAs were used to calculate the interaction of two variables of the simulation activity and the level of expertise. To evaluate if the differences in performance was due a learned effect over the course of the simulation session, a two-way mixed ANOVA with repeat measures was performed. Statistical analysis was performed using SPSS and Laerd Statistics (2015). Significance was assessed at $p<0.05$.

Qualitative Data Analysis

The transcripts from the simulation events were coded with codes modified (figure 5) from Crowley, Naus, Stewart and Friedman (2003) who investigated visual diagnostic

performance expertise in pathology. Crowley and colleagues (2003) used large categories of data examination, data exploration, interpretation, control process and operational verbalizations. Using this as a foundation, the action codes were divided based on anatomical identification (AC1-AC11), localization and reasoning (AC12-AC14), strategies for matching (AC15-AC18), and experience and confidence cues (AC19-AC30) (Figure 5). This coding scheme closely followed the different steps being asked of the participants during the simulation events. The coding provided a method for comparing the reasoning during the simulation between experts and novices.

To evaluate the reliability of the coding scheme developed for this study, ten percent of the data was given to another researcher to code after initial instruction. According to Cohen's κ, there is good agreement between the two researchers' judgments, $\kappa = 0.619$, 95% CI [0.574, 0.664], $p < .0005$. Good agreement for κ is between 0.61 and 0.80. However, there were codes that required heavy medical anatomy understanding. The research coder did not have expertise in medical anatomy, so those codes were removed from the Cohen's κ calculation. A Cohen's κ for this data is slightly higher and suggests good agreement between the two researchers' judgments, $\kappa = 0.674$, 95% CI [0.627, 0.721], $p < .0005$. The first author coded the remaining 90% of the transcripts.

The transcripts of the test case and the following six simulation events were coded. There were 77 coded events for the novices, and 68 codes events for the experts. The transcript for each event was coded for the presence of the overarching action codes (Figure 5). Each time an action code was used in that event, it was given a score of '1'. That action code could be used multiple times in an event. The action codes were summed over all 77 events for novices and 68 events for experts. A percentage of times that action code

was present were obtained by dividing the total action code number was by the number of events (77-novices or 68-experts). The percentage (proportion) of these codes used by novices and experts were compared using the "N-1" Chi-squared test (Campbell, 2007; Richardson, 2011). Statistical analysis was performed using SPSS and Laerd Statistics (2015). Significance was assessed at $p<0.05$.

This analysis looked at each independent action code and compared them across the novice-expert. However, it is likely that some of these action codes are related and analysis with that lens may lead to less significance, however clusters of codes may provide valuable insight into specific areas of competency that may not be reflected in individual action code assessment.

Results

The qualitative and quantitative results of this study indicate differences between experts and novices along several of Bransford and colleagues (2000) dimensions. I will use those dimensions to organize the findings below.

Flexible retrieval with little effort

Experts can more flexibly retrieve their knowledge with little conscious effort, (Azevedo, 1999) and frequently perform these tasks faster than novices (Chi, 2006a; Crowley, et al., 2003; Gould, 1978; Stevens, et al., 2010). This study corroborates these findings (table 1), as assessed through both time and number of probe movements when matching up the target image to the location in the simulation torso.

Experts were faster than novices independent of imaging plane presented for the event. In the axial plane, experts' average time from ready to completion was 16.6 sec compared to 26.6 sec for novices (two-way ANOVA; F=6.232, p=0.014). Similarly in the sagittal plane, experts' average time was 19.99 sec compared to 38.6 sec for novices (F=10.693, p=0.001).

Experts were equally proficient in performing the task in both the axial and sagittal orientation (F=0.441, p=0.508). However, novices' were significantly faster with axial orientated CT images compared to sagittal CT images (F=6.6, p=0.011). Sagittal images are less commonly used and localization of the anatomical structures and processing of the image is representative of the right or left side of the body uses more advanced retrieval and application of anatomical information. Images in the sagittal plane add an order of complexity to the imaging that most novices may not have grasped. Thus it is not surprising that experts process these images better than novices.

In this study, the time to localize the image is likely conflated with the think aloud process. One possible reason for experts being faster than novices might be that novices talked more during the sessions (novices 83.3 % compared to experts at 70% of overall events). However, there was no statistical difference in the number of sessions that the novices and experts verbalized (χ^2= 0.287; p=0.998), thus this does not appear to be a factor affecting speed.

In addition to experts being significantly faster in all components of the simulation components, (table 1) (ready to completion and ready to visual threshold), they also perform the task of matching the target image with their localized image with significantly less movements of the probe, 36 movements compared to 61 movements for novices (one

way ANOVA $F=14.948$; $p<0.05$). Experts are more intentional and direct compared to novices. This may also reflect a confidence in knowing the anatomy and the direction of movement needed to get to the correct anatomical location. The time and movement data indicates that experts do possess more flexible retrieval than novices when it comes to coordinating radiological images and physical anatomy.

There is also a consideration that there is an improvement in event time from the first event to the last that could account for differences in event times. A two-way mixed ANOVA with repeat measures analysis was performed that determined that the event order was not significant ($p=0.422$) (Durbin-Watson = 1.613), but rather expertise was a greater factor in the event time.

Content knowledge structured and organization

Expert knowledge is organized around core concepts or big ideas, and is not a list of facts or formulas within a domain (Bransford, Brown & Cocking, 2000; Glaser, 1984). I find evidence that experts in my study organize their knowledge by relating to prior experience, using external landmarks to describe the location and defined the spatial anatomy.

Experts tended to structure their content knowledge around core concepts that make sense from a perspective of clinical experience; they approached each event as a clinical problem. Although both novices and experts used references to physical landmarks to isolate the internal anatomy, there was a (non-significant) tendency for experts to use these references more ($p= 0.2309$; Table 2). Experts were more likely to relate to past experience and how imaging is different between patients thus rely on general anatomical principles when interpreting imaging exams rather than the specific image details (like

novices). They did not take the images at face value, but related it back to past experience and how they would define this spatial relationship in the body.

Novices on the other hand, took each case in literal context. They keyed into identifying individual structures ($p=0.0338$; Table 2), and the size and shape of organs ($p<0.0001$; Table 2). In doing so, novices are more likely to misinterpret the anatomy ($p=0.0021$, Table 5) or have uncertainty of the anatomy that they are viewing ($p< 0.0001$, Table 5). A common example was the misinterpretation of the third segment of the duodenum and referring to it as the pancreas. Another example was the misinterpretation of the lumbar spine on an axial CT image, and referring to it as the thoracic spine. The novice did not notice the size and shape of the lumbar vertebral bodies, and the lack of ribs. Additional errors in anatomy were identifying structures that were not in the image (i.e. the presence of the pancreas or renal arteries). These examples, demonstrate that the novices knowledge was organized around details rather than the larger concepts. By focusing on the individual details, they did not see the associations of the surrounding anatomy to help distinguish the pancreas from the duodenum, or the other anatomical features to differentiate the lumbar from thoracic spine.

Novices used non-anatomic descriptive cues (color, blobs, patterns) to describe what they are viewing ($p=0.0042$, Table 2). Some examples of these are using the pattern of the intestines to localize the image. Many of the experts voiced that this was not a reliable way to evaluate the abdomen, as the intestinal pattern is highly variable within and between patients. Novices would also use the air bubbles in the intestine to match up. They used abstract black/white (gas in the intestine, 'brightness of the bones') over anatomical structures.

Novices organized their knowledge around non-anatomical idiosyncratic things whereas experts organize around the regional anatomy presented and their prior clinical experience. Experts generalized their anatomical reasoning across all variations of clinical presentation, and novices concentrated on the literal presentation of the image. The novices would go through a 'list' of structures more often than concentrating on the big concept of where the image was located.

Recognize complex patterns

Experts are more likely to recognize meaningful configurations and realize the implications (Bransford, et al., 2000). Experts have the ability to chunk information into meaningful and familiar patterns, and apply this to a situation. I find evidence the experts in this study demonstrate this ability in recognizing organ substructures, using small structures for fine tuning, and using structural differences in the surrounding anatomy.

Expert's notice features and patterns that are not noticed by novices. When novices were focused on the presence of the larger structures (liver, kidney) the experts were identifying structures that would be more helpful in making an accurate localization of the image. Experts noted those larger structures but would drill it down further. When localizing an image in the body, experts rely on organ substructures (experts 69%, novices 19%; $p<0.0001$, Table 2; AC4) whereas novices weigh heavily on size or amount of an organ in the image (Experts 71%, novices 127%, $p<0.0001$, Table 2; AC7).

Demonstrating their ability to recognize meaningful complex patterns, experts significantly used small structures for fine-tuning of matching the images (Experts 75%, novice 27%; $p<0.0001$, Table 4; AC15). Experts would use smaller anatomical findings to

fine-tune their search such as the consistent narrowing of the slope of the lungs with inferior scanning or the isolation of a vessel in the abdomen, which will have a narrow zone of transition in the given plane. They narrow the complex problem by isolating a small and reliable structure that will have a dependable change. In contrast, during this same task, novices typically rely on matching at the organ level (size, shape) (Expert 78%, novice 136%; $p<0.0001$, Table 4; AC16), use non-anatomic cues (Expert 21%, novice 61%; $p<0.0001$, Table 4; AC17) or use trial and error methods (Experts 1%, Novice 29%; $p<0.0001$, Table 4; AC18).

To provide a rich sense of these statistical effects, consider some extended qualitative excerpts from the think alouds. When localizing an image in the upper abdomen (Image KC1, Figure 4) that does contain the liver, novices concentrate on the presence of the organs. Here is an example from one novice interview.

> 'So I see the liver here. Um, and it's quite large; so I know it's probably a little bit near the diaphragm. I think this is the spleen right here, and I think that's the pancreas. So I'm going to go where I think that all of that stuff is….. I'm going to go down a little bit because I think I'm getting some of the like diaphragm here to where the spleen is a little bit larger……. I was focusing in on like the smaller organ and the spleen because I felt like I could match that up better than using the liver which is larger. So it's the size of the spleen is what I was using on that one. And I mean, the shape. Obviously if like the shape changed drastically, I think I would go the other way if I thought. But mostly the size'

For the localization of this same image, this expert notes the dome of the liver, but uses the sharp change in the costophrenic shape of the lung as a more accurate landmark.

> 'when I was trying to localize, I was looking the dome of the liver, but when I'm trying to match the images, then the amount of costophrenic recess I have….. the black space of the costophrenic recess is an easier -- because it's an AP dimension, I can just kind of go with that. If I'm doing the liver, you know, every liver is a little different, and so depending on how much like of the left hepatic lobe I'm including. I think it's a little bit -- it changes less than the size of a costophrenic recess when you get that low.'

In this example the novice uses trial and error and even incorrect anatomical identification in their localization (there was no pancreas in this section). In contrast, experts isolate important landmarks that define the space and more specific cues to further define it. Expert's demonstrated the ability to recognize meaningful patterns, whereas these patterns did not come to awareness for the novices.

A similar scenario occurred with the anatomical structures in the middle of the abdomen (Image KC14, Figure 4). The novice concentrated on the large structures and also based judgment on a structure that greatly changes over time (intestinal air).

> 'I see kidneys again, and then I see the bottom of the liver, what looks to me like small bowel, though I'm not great at distinguishing bowel……So I notice a little bit of difference in the bowel pattern, but otherwise I see kidney and the tip of the liver, so I'm just going try to match up the bowel pattern. But specifically like the air --'

An expert given this same image, concentrated on the renal vessels that have a horizontal orientation and a narrow zone of transition.

> ' -- so I'm at the inferior margin of the ribs and the kidneys. So again I'm going to be about -- I'm above the umbilicus. …..-- so I'm looking at the right renal vein -- almost. And so I'm just trying to line that renal vein up there.'

This example demonstrates the novice using unpredictable landmarks (intestinal air) and large structures to localize, compared to experts identifying a structure with a narrow transition to solve the complex problem of localizing in the mid abdominal region. The expert was able to recognize the relationship and organization of the anatomy to do the final localization, rather than doing a search of what is in the region. The expert had a better defined plan of execution.

Exhibits situational applicability

Experts can retrieve the knowledge appropriate for the task. Albeit this simulation session was not in the usual radiology reading room context, experts were able to apply the appropriate terminology and anatomy of the CT imaging to a simulation tool. Experts are more likely than novices to use the correct terminology ($p<0.001$) (Table 2). Novices tended to use lay terms to describe anatomical structures, such as belly button for umbilicus, or 'bowel' for large intestine or small intestine, or references like 'hipbone', 'front end of the stomach', 'liver shadow', or 'stomach bubble'.

'Stomach bubble' or 'liver shadow' would not be terms that experts would use in describing findings or organs on CT images. The experts stayed in context to what medical imaging study they were using and not incorrectly interchanging terminology. All the

images provided in this exercise were CT images; ‘stomach bubble’ is not a CT descriptor, yet it was a term used by novices. The term, ‘stomach bubble’, can be used on plane films of the chest or abdomen to identify the gas within the stomach. The reference to liver shadow is more appropriate to ultrasound imaging when image characteristics can be described using shadowing of the sound waves.

Experts know their discipline thoroughly and can apply knowledge to various situations appropriately. The novices may have learned these terms in their prior courses and clinical experiences, but they might not have had the time or practice to connect those terms to the appropriate situation (in this case the CT scan).

Easily adapt to a new situation

Experts can ‘act in the moment’. Experts do this on a daily basis with each case and patient that they encounter. They approach new problems to use their expertise and look at it as an opportunity to expand their current level of expertise. Doing so requires, confidence that they can adapt to and succeed in a new situation.

During the study, participants commented about their level of confidence or their ability to incorporate past experience into their reasoning through each event. On average, experts made more confident statements ($p<0.0001$, Table 5) and incorporated past experience ($p=0.0012$, Table 5) more often than novices. For example, experts made statements such as, "It's experience now, I mean -- ", 'I am not sure how I know, but I just know", "It's a pattern I have in my head. So I suppose it is a concept that I've developed over years, you know ", and "from experience looking at lots, thousands of CTs of the belly."

On the contrary, novices significantly expressed uncertainty in their knowledge (Experts 18%, Novices 101%; $p<0.0001$, Table 5, AC24) and were more likely to verbalize that the task was difficult for them ($P=0.0378$, Table 5, AC29). Common statements were: "Then I guess here", "I'm not very familiar with this one", and "I really don't know". Their verbalization of difficulty was at times more specific: "Sagittal is just much more difficult for me", "I feel like we see the most axials so then we have to like switch everything up [referring to looking at sagittal images] to think about the anatomy, It's hard" and "I don't have a good sense of the 3D - how the pelvis would change, so I'm like having to do little testing for that". These quotes and the uncertainty in their knowledge demonstrate that novices have trouble adapting to new situations, while experts do not.

Discussion

Given the differences in how experts and novices correlate radiology images with the physical body, how can we better support medical students in achieving expertise? Ericsson proposed that expertise is a result of 'deliberate practice', where the learner has full mental engagement that is goal oriented and focused on overcoming current performance boundaries (Ericsson, 1996). He proposed that expertise is a product of 10,000 hours of deliberate practice (Ericsson, 2008). Does this type of deliberate practice that supports the development of the features of expertise occur in medical education? If not, what changes could we make such that it does?

Current state of medical school education

The current medical school curriculum, as two years of basic science followed by two years of clinical experience, introduces students to topics, but does not afford students the opportunity to go into depth. There is no mandatory requirement for research investigation. There is limited cross collaboration with other courses, which later impedes integration of the material in clinical medicine. The principle understanding of medical school is to introduce students to topics, and if they are interested, they can go deeper on their own. Medical school thus acts a primer for the apprenticeship model of medical residency programs. Following residency, many residents are compelled to pursue fellowships to gain more extensive training not presented in their residency.

Medical students in the first two years are predominantly assessed with multiple-choice question (MCQ) exams. No fill in the blank or short answer. The MCQs provide convenient standardization and efficient testing for large classes (Vanderbilt, Feldman & Wood, 2013). MCQs are often poorly written, and test memory recall of individual facts, rather than knowledge application (Vanderbilt, et al., 2013). These types of tests do not assess critical thinking skills.

The current curriculum and methods of assessment in medical school does not adequately prepare students to become experts. The topics in medical school are presented to expose students to many disciplines, however this does not allow enough exposure for the students to self-organize the content meaningful 'chunks', which will later allow for more efficient retrieval. Without the time to organize their knowledge into meaningful chunks, the students are at a disadvantage to be able to recognize complex patterns when presented. Because medical students haven't had the opportunity to

develop a deeper knowledge for specific content, it is unlikely that they can exhibit 'conditionalized' knowledge. This lack of deep understanding also impedes the medical student from being able to flexibility retrieve structured knowledge with little conscious effort in critical situations. Lastly, the structure of the curriculum and assessment methods does not prepare the students to 'act in the moment', and apply their knowledge to new situations. It does not promote critical thinking that fosters this behavior.

Possibilities for Improving Medical Education

This study again demonstrated that experts could retrieve knowledge with little effort and also do this faster than novices. This ability did not come without years of practice and developing their knowledge content and organization. Today's students have grown up surrounded by technology, and it is part of their daily lives. Technology is great at storing information for later retrieval (Ferguson, 2010); however, we need to foster long-tern learning and retention in students that is independent of technology. We need to encourage learners to function independently, and build upon intrinsic motivation and natural curiosity (Ferguson, 2010). This requires learning activities that apply higher-order thinking skills, and activities that mirror real world situations (Ferguson, 2010). Through these activities the students will increase their depth of knowledge and interconnections to be able to retrieve this knowledge without effort. Within radiology, medical students could be encouraged to look at images of patients and get use to using the PACS (picture archiving and communication system) workstations and try to interpret the images first and discuss the findings with the radiology faculty.

Experts organized their knowledge around core concepts that make sense from a clinical experience, treating each event as a clinical problem. In contrast, novices took each case in literal context, listing findings and even concentrating on details that were non-anatomical. Often exams in medical school exams focus too much on the recall of independent facts rather than the application of knowledge (Vanderbilt, et al., 2013). This behavior translates into students concentrating on details, rather than learning critical problem solving skills. To aid students to develop the skills of experts, curriculum should focus on content but also problem solving using clinical cases. Critical thinking skills should be integrated into core concepts that are supported by the details, and illustrates the importance of these details to the core concepts. The assessment of this content should focus on application of this knowledge utilizing critical thinking skills. Within radiology, students could not only identify the salient features within an imaging exam, but also discuss a differential diagnosis and what imaging features support those diagnoses.

Experts are able to notice features and patterns that are not noticed by novices. They are able to chunk information to meaningful packets and apply to a situation. Novices used large structures to localize, used non-anatomic cues or resorted to trial and error methods. It is important to point out these salient points in a subject to students, however it goes beyond that. We should be showing them how to use these salient points, and also encourage them to search for them in other disciplines. These features and patterns are not only in radiology, but also in every discipline of medicine. Showing learners how to identify these patterns, and encouraging them to look for such patterns, will promote responsibility for their own learning. In radiology, students can work one-on-one through cases with radiologists and discuss these key findings. Later, this understanding on how to

use these features and patterns can be evaluated by students demonstrating how they used these patterns in unknown cases.

Experts can retrieve the knowledge appropriate for the task and can 'act in the moment'. They approach new problems to use their expertise and look at it as an opportunity to expand their current level of expertise. Doing so requires, confidence that they can adapt to and succeed in a new situation. The novices in this study were not always comfortable using the correct terminology or lacked confidence in applying their knowledge to a new of difficult situation. In medical school we need to encourage the students not to think like students but to think and act as professionals. We need to encourage students to think, make decisions and solving problems through integration and assimilation of knowledge (Finnerty, et al., 2010; Mann, 2010). There are many ways to accomplish this goal, but one way may be through simulation exercises that create real world situations that force students to apply their knowledge and think in the moment (Finnerty, et al., 2010). Simulation sessions also promote practicing and reinforcing experimental inquiry (Finnerty, et al., 2010), which will lead to confidence when called to 'act in the moment'. There will be failures in a simulation exercise, but it is an ideal method to learn areas of knowledge gaps with no detrimental consequence in a true clinical situation.

Conclusion

Experts and novices in radiology have similar features to experts and novices in other domains - inside and outside of medicine. The results from this simulation study on novice-expert differences demonstrate that experts are faster, use medical terminology

appropriate to the specific imaging, identify cues (narrow zone of transition) to fine tune image localization, and recognize meaningful patterns to focus on an anatomical region. In contrast novices were more likely to identify anatomical structures in a list fashion, concentrate on organ level anatomy and compare locations by organ size and shape.

The analysis evaluated each action code as an independent event. However, it is highly likely that there are meaningful cluster of codes that could be analyzed and provide valuable insight into competency that is not reflected in the individual action code assessment.

The expert-novice paradigm is useful for understanding these major differences – and this study has helped us identify the starting and ending points in learning to coordinate physical exam with imaging. However, there are likely many steps along the way – and further work is needed to understand the trajectory from expert to novice. One method would be to follow Dreyfus and Dreyfus' (1980, 2004) model of adult skill acquisition and study learners moves beyond the expert-novice dichotomy and instead looks at five (rather than two) stages of development. These five stages can be applied to medical education, from medical students to experienced clinicians. More investigation is needed to define these stages (novice, competent, proficient, expertise, master) within medicine and within a specific medical specialty. A deeper understanding of the knowledge and skills at each level can provide specific guidance for defining appropriate curriculum and evaluation processes for medical students and residents. The novice-expert research paradigm opens up targeted opportunities to make our medical education system stronger.

Figures

Figure 1: Components of simulation tool. A, iPad image viewer and control panel; B, handheld probe that moves image set in integration with the torso; C, torso model for physical integration with the image set and handheld probe; and D, mounted box housing all the controls.

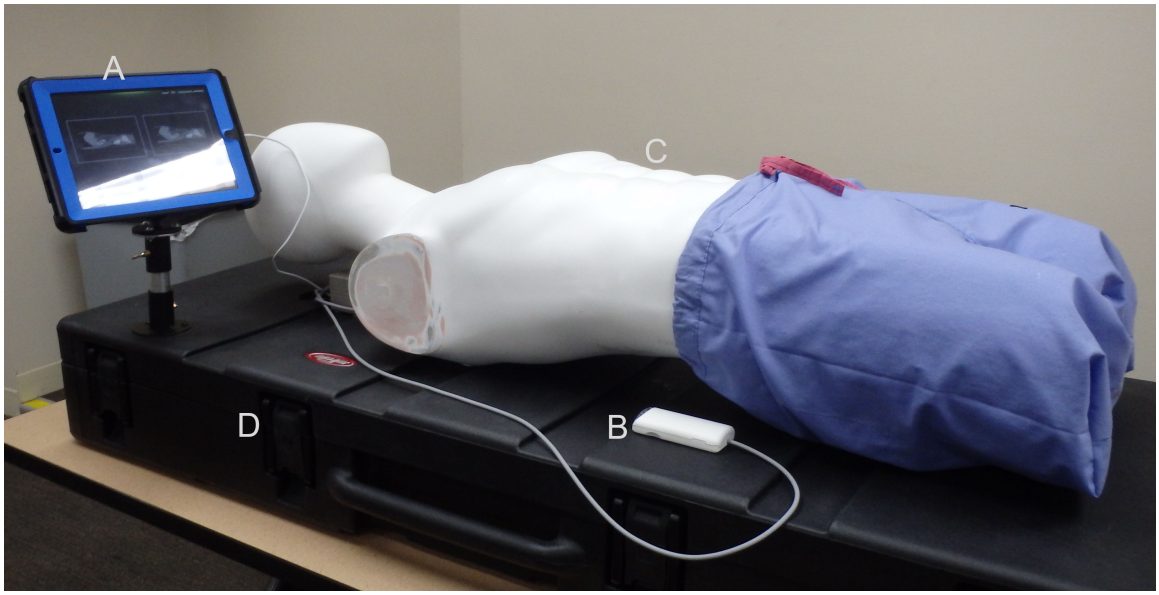


Figure 2: Interaction of the participant with the simulation tool localizing an image within the physical torso

Figure 3: Location in the simulation torso where the image slices are located (axial (pink tabs) and sagittal (yellow tabs), purple tab represents test location)

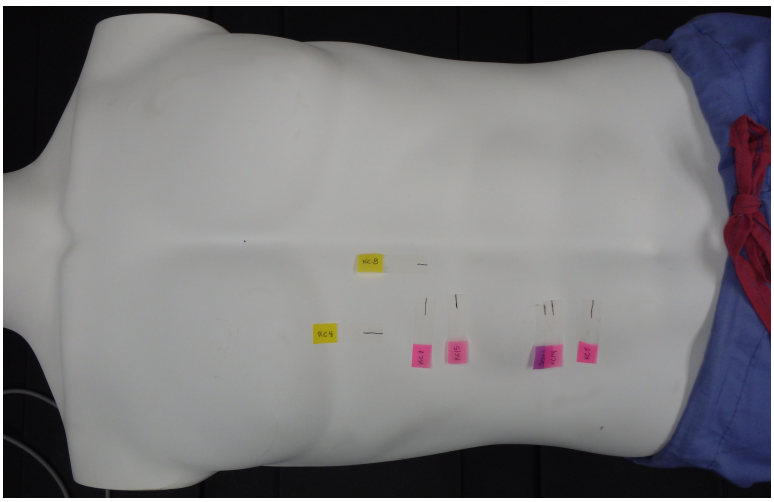


Figure 4: Image presets for participants to localize (from left to right and top to bottom; KC15, KC14, KC8, KC1, KC5, KC4)

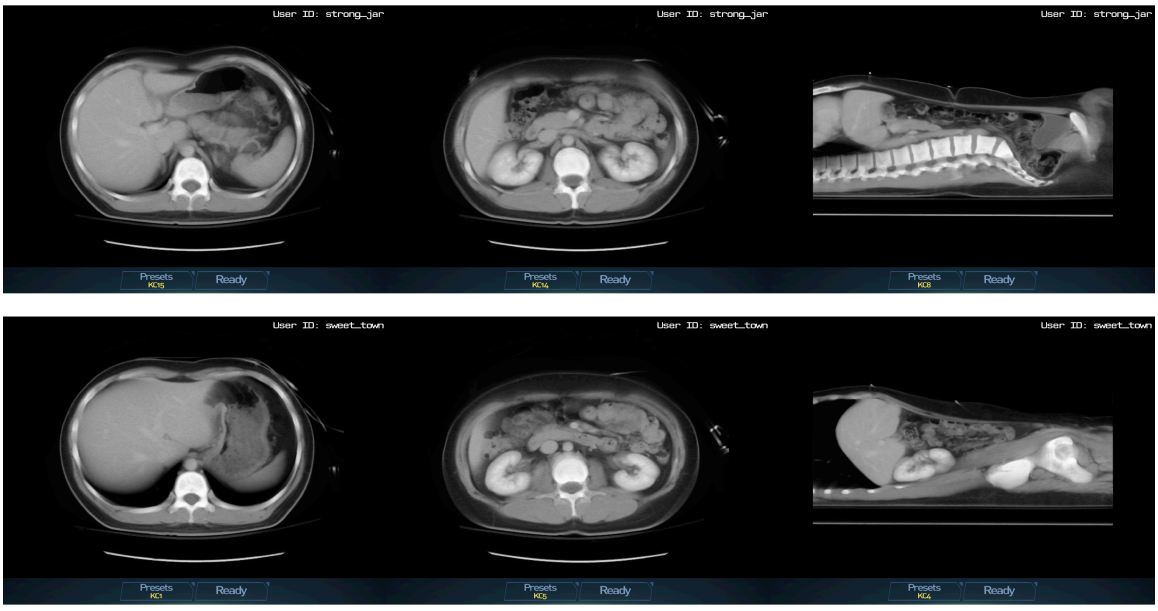


Figure 5: Coding schemata. Modified from Crowley, Naus, Stewart and Friedman (2003)

Number	Action Code	Description of code	Examples
AC1	Identify imaging plane	Describes the imaging plain of the presented image	"So this means another axial image"; "so we're sagittal'
AC2	Identify anatomical structures	Describe what they are seeing in the image at the time of the initial assessment; Names the organs that they see in the image	"stomach, spleen", "small bowel, and bottom tip of the liver"; "spine, vertebrae, kidneys"; "left renal vein"
AC3	Non-verbalization	No dialogue while doing task	Silence; often noted in transcripts as pauses or lapsed time
AC4	Emphasizes substructures of organs	The identification of these substructures is important to their understanding of the image location	"I can see the rugae in the stomach"; "medial limb of the right adrenal gland"; "third segment of the duodenum crossing over."
AC5	Physical contact with the mannequin	This actions simulates action similar to a physical examination	Palpation of the mannequin to localize or define the location of the anatomy, similar to a physical examination
AC6	Uses an external landmark to describe location	Uses landmarks on the external surface of the body to describe the anatomic location	"It's above the umbilicus"; "I want to be near the epigastric region"; "mid-clavicular"; "use bottom rib to signify where liver edge is located"; "the in-bowing of the -- of the chest wall, so I'm below the Xiphoid but not that far"
AC7	Shape, size or amount of organ	Describes the size or shape of the organ or structure. It becomes of some importance to identifying the organ location	"so the liver is quite a bit larger"; "larger slice of liver"; "liver took up more of the screen"
AC8	Expresses the importance of adjacent anatomy	Uses the adjacent anatomy to describe the location of the anatomy (presence or absence of anatomy/organs). This relationship of the anatomy is important in their	"So level of like the lowest ribs with -- you can still see part of the liver"; "And so, they're just, like, easy to see how they related to the anatomy around them"; "And then with my adjustments on last time, I know the kidneys are south of the liver -- "

		localization.	
AC9	Location within an organ	Describes a non-anatomic terminology location within an organ to localize where the image is in the organ	"biggest part of the liver"; "upper like third-ish or like the lower third (when describing a kidney)"; "superior/inferior pole of the kidney"; "lower part of the liver"
AC10	Use of incorrect terminology	Describes anatomic findings using non-medical or non-anatomic terms	"stomach bubble"; 'air fields" (when describing lungs); 'Belly button"' "bunch of bowel"; 'big lung window"; "gastric bubble"; north and south rather than superior and inferior; describing structures by color, BW, contrast rather than the organ.
AC11	Use of descriptive cues (non-anatomic)	Describe structure by its color	" And then there are these three little dots in the duodenum, and so I matched up those"; "I was seeing if …this little, black blob - if that's similar to that one"; "liver shadowing."; "This time, I think it's more contrast in color."
AC12	Identify Location within the body (initial)	Describes the location of where organs/structures are in the body	"this is pretty high, I see the stomach"; this is on the right- hand side"; " Well I knew it was right, because his liver was there"; "its above the umbilicus"
AC13	Reasoning why used certain organs/ structures to localize (initially)	Using key organs or structures to identify the location of the image	"So I'm well medial to the iliac crest. So again that would bring it in here rather than being out here"; "And the pancreas is up pretty high, so I would say, like, I don't know -- there-ish?"; "The relationship, yeah, because knowing the psoas I know that it has to be a little more medial because if I go too far lateral, it's going to be the quadratus lumborum because I've come out of the psoas. "

AC14	Compare to expected cued target	Compare to the target goal and where they placed the probe.	"I'm too high", "I'm too low"; "I'm too far to the right"
AC15	Elements or reasoning for fine tuning using small structures (matching up image)	Using structures that have **narrow zones** (vessels, loop of bowel) of change to adjust the location and match images	"Well, it's thinner in the superior/inferior dimension. So it's hard to -- things that are thicker are harder to fine tune like that";" Because it was the tiniest structure that I could use that was precise, like as opposed to like a triangular piece of the liver or, you know, kidneys you can't really, you know, they're kind of look the same for a period of time, and so I was using the renal vein"; "Yeah, you want - vascular structures are going to be helpful. You see that that's what I've been going after, because they're relatively unchanging"; "I was just looking at the way the crescents (of the lungs) looked. "; "so I was looking at the loop of small bowel that goes across the middle of the image"; "looking at the shape of the bone here and seeing what, you know, sort of the shape of the iliac crest"
AC16	Elements for reasoning remains at organ level (matching up images)	Using the size and shape of a larger organ (liver, spleen, kidney) to adjust location	"I was focusing in on like the smaller organ and the spleen because I felt like I could match that up better than using the liver which is larger. So it's the size of the spleen is what I was using on that one."
AC17	Reasoning for matching up that is not anatomic	Using non-anatomic references to match up images; matches colors or patterns	"The air and the fluid level in the stomach is what I was matching it up on"; " looking for patterns, like seeing this kind of swirl part of the kidney or looking for that dark spot there that looks similar to the test image or the initial image and looking for the bowel gas patterns"

AC18	Trial and error for matching and refining the matching of anatomy	Makes movements in search of specific anatomy	"going to keep moving down until I see the kidneys"; "And then I was, like, getting positive reinforcement from the image as I was moving down that it was adjusting more to the correct one. "
AC19	Uncertainty in anatomy	Uncertain what structure they are seeing in the image	"looks like a large vessel. I don't know if that's the aorta or not, but pretty close to midline. Puts me on the left side. Slightly in the left side. But it could be the IVC, in which case I would be on the right and I really don't know."" I'm not sure what that black is (in reference to the lungs)";"I think that's the pancreas"
AC20	Mis-interpretation of anatomy	Called the presented anatomy the wrong structure; don't know what the structure is; not knowing where anatomy is on mannequin (internal or external)	Referring to anatomy that is not present, i.e.. I think that's the pancreas (there is not pancreas in the image); Referring to the duodenum as the pancreas; referred to anatomy in the image that was not present in the image (i.e. Gall bladder's presence when it was not present); "He doesn't have a belly button" (but it really does have one)
AC21	Compare at case level/ prior cases	Compares the current image to a prior case - either appearing similar or in relationship to another case within the simulation session	"I'm just a little higher than the last one"; "So I'm going to go a little bit lower than I did the last time"; "this is similar to the first one"; "this is like where I had initially put my probe because I saw too much of the liver"; "So I used the basis of what I had to move the last time"; "I know it's higher that the last one"; "Kind of looks like the first one you showed me"
AC22	Localizes by the absence of a structure	The absence of a structure becomes an important way of localization	" I know that I'm on the right side because I don't see the heart"; " so we're not seeing much spine"; "I don't see the kidneys yet"; "I'm not midline because I don't see the vertebrae"

AC23	Certainty of knowledge (positive)	Statements about their confidence in what they are doing (affirmative statements; sure about their decision)	"It's experience now, I mean -- "; "I am not sure how I know, but I just know"; "I'm correct"
AC24	Uncertainty of knowledge (negative)	Statements about their lack of confidence in what they are doing (uncertainty; not completely sure)	"Then I guess here"; "I'm not used to really trying to determine where on the body the pancreas is"; "I guess"; "I'm not very familiar with this one"; "I really don't know"
AC25	Direction to anatomy	**Vocalization of direction to move** to locate desired anatomy (correction to initial localization)	"we're down at the very edge of the liver, kidneys so I got to go farther inferior"; "So I need to go move towards where there's more volume of the liver."; "The liver was too big, so I knew it was too high."
AC26	Incorporation of past experience	Relates how they approach the case using reference from their own past experience outside of this simulation event	"It's a pattern I have in my head. So I suppose it is a concept that I've developed over years, you know"; "from experience looking at lots, thousands of CTs of the belly"; "No. I'm afraid now I think CT. I always, even in ultrasound, start with the axial plane because I think CT"; "Like this reminds me of when like you scroll down on the CT chest too low and you start seeing the liver."
AC27	Define spatial anatomy for themselves	Their explanation of how they are able to localize the image in space	"I have a very good 3-D picture in my head of where it's supposed to be"; "So I visualize where it should be physically in space and where I have to interact with the body's surface to get there"; "Yeah, I'm just picturing the anatomy of the organs in the abdomen. So, yeah"; " I don't have a good sense of the 3D how the three pelvis would change, so I'm like having to do little testing for that. "
AC28	Unreliable anatomy	Explains why some anatomy is less reliable to use for localizing	"between studies, they vary so much because you have peristalsis going on all the time. So bowel is really…..bowel is too variable."

AC29	Verbalizes difficulty	When localization of the area or plane of imaging is more difficult to assess	"Sagittal is always more difficult"; "Yep. And actually midline is the hardest anyways"; "I don't have a good sense of the 3D how the three pelvis would change, so I'm like having to do little testing for that. "
AC30	Realization of anatomic cues that they missed	Retrospect realization that there was key anatomy that could have helped guide their anatomic location	"I guess and the belly button -- you see exactly the belly button. [CHUCKLING] That one should have helped me"; "Well, no. I should have seen that. I didn't look at the umbilicus. That was stupid"; "Again, you got to look at all pictures of the image."

Tables

Table 1: Time comparisons between novices and experts.

Event Interval	Expert Average Time (seconds)	Novice Average Time (seconds)	F-statistic*	Significance*
Pretest to ready	20.48	44.11	54.741	$p<0.05$
Ready to completion	17.76	30.57	14.978	$p<0.05$
Preset to visual threshold	29.2	57.5	46.572	$p<0.001$
Ready to visual threshold	8.7	13.4	6.585	$p<0.05$

*One-way ANOVA with $F(1,121)$ using SPSS.

Table 2: Collated Data Exploration of Novices and Experts

No.		% events with action code		Chi-squared Analysis				
	Action Code	**Expert**	**Novice**	**Difference**	**95% CI**	**X^2**	**DF**	**p-value**
AC1	Identify imaging plane	15	16	1%	-12.02 to 13.599	0.027	1	0.8687
AC2	Identify anatomical structures	232	309	77%	-170.6 to 324.6	-4.507	1	0.0338
AC3	Non-verbalization	37	27	10%	-6.07 to 25.75	1.657	1	0.1981
AC4	Emphasizes substructures of organs	69	19	50%	33.74 to 63.36	36.698	1	<0.0001
AC5	Physical contact with the mannequin	6	17	11%	-0.384 to 22.103	4.194	1	0.0406
AC6	Uses an external landmark to describe location	54	44	10%	-7.19 to 26.60	1.435	1	0.2309
AC7	Shape, size or amount of organ	71	127	56%	26.55 to 86.08	-1512.841	1	<0.0001
AC8	Expresses the importance of adjacent anatomy	15	27	12%	-2.33 to 25.59	3.073	1	0.0796
AC9	Location within an organ	57	51	6%	-11.098 to 22.69	0.519	1	0.4711
AC10	Use of incorrect terminology	13	60	47%	31.298 to 60.00	33.638	1	<0.0001
AC11	Use of descriptive cues (non-anatomic) - "color, blobs, patterns"	12	32	20%	5.62 to 33.367	8.195	1	0.0042

Table 3: Reasoning of how image relates to body initially and compared to target

No.		% events with action code		Chi-squared Analysis				
	Action Code	**Expert**	**Novice**	**Difference**	**95% CI**	**X^2**	**DF**	**p-value**
AC12	Identify location within the body (initial)	101	123	22%	-2.166 to 45.794	-12.145	1	0.0005
AC13	Reason why certain organs/ structures were used in localization (initial)	60	69	9%	-7.45 to 25.12	1.273	1	0.2592
AC14	Compare to expected cue target	79	81	2%	-11.82 to 16.10	0.09	1	0.7644

Table 4: Strategies for fine-tuning of anatomical structures

No.		% events with action code		Chi-squared Analysis				
	Action Code	**Expert**	**Novice**	**Difference**	**95% CI**	**X^2**	**DF**	**p-value**
AC15	Elements or reasoning for fine tuning using small structures (matching up images)	75	27	48%	31.51 to 61.58	33.053	1	<0.0001
AC16	Elements for reasoning remains at organ level (matching up images)	78	136	58%	20.38 to 96.26	-126	1	<0.0001
AC17	Reasoning for matching up that is not anatomic	21	61	40%	23.471 to 54.11	23.518	1	<0.0001
AC18	Trial and error for matching and refining the matching of anatomy	1	29	28%	16.45 to 39.50	21.059	1	<0.0001

Figure 5: Confidence in knowledge and judgments of Experts and Novices as they correlate physical examination and cross sectional imaging

No.		% events with action code		Chi-squared Analysis				
	Action Code	**Expert**	**Novice**	**Difference**	**95% CI**	**X^2**	**DF**	**p-value**
AC19	Uncertainty in anatomy	9	55	46%	30.91 to 58.66	34.099	1	<0.0001
AC20	Misinterpretation of anatomy	10	31	21%	7.03 to 34.007	9.483	1	0.0021
AC21	Compare at case level/ prior cases	38	40	2%	-14.72 to 18.48	0.06	1	0.8061
AC22	Localizes by the absence of a structure	16	27	11%	-3.45 to 24.74	2.542	1	0.1109
AC23	Certainty of knowledge (positive)	88	43	45%	29.44 to 58.04	31.559	1	<0.0001
AC24	Uncertainty of knowledge (negative)	18	101	83%	71.38 to 92.48	104.941	1	<0.0001
AC25	Direction to anatomy	56	75	19%	2.58 to 34.56	5.777	1	0.0162
AC26	Incorporation of past experience	62	35	27%	9.80 to 42.59	10.48	1	0.0012
AC27	Define spatial anatomy for themselves	51	32	19%	1.996 to 34.991	5.355	1	0.0207
AC28	Unreliable anatomy	15	10	5%	-6.69 to 17.14	0.829	1	0.3627
AC29	Verbalizes difficulty	10	23	13%	-0.116 to 25.5	4.315	1	0.0378
AC30	Realization of anatomic cues that they missed	10	9	1%	-9.56 to 12.07	0.042	1	0.8379

Chapter 4

Novice-Expert evaluations of peer-based medical imaging diagnoses

Abstract

PURPOSE

Medical care depends upon the ability to reason through data and evaluations from other practitioners. Often this process necessitates the evaluation of a peer's conclusion. The ability to make these evaluations is an important responsibility of medical practice and patient care. Within the novice-expert research paradigm, this study will look closer at novices and experts use of their knowledge of recognition and recall to solve medical imaging problems.

METHODS

Ten radiologist experts (average age 47.4 years with 13.5 average years of experience; 9 males, 1 female) and 11 senior medical student novices (average age 27.4 years; 5 males, 6 females) performed a simulation using axial computed tomography images within a human simulation torso. This study was IRB approved. Quantitative data were collected on the evaluation of peer–assigned image localization and normal/abnormal judgments of anatomical structures. Participants engaged in qualitative "think alouds" about their reasoning. The transcripts were coded and assessed for emerging themes during the evaluation of the biopsy approach. Chi-square analysis was performed on the qualitative codes. Significance was assessed at $p < 0.05$.

RESULTS

Experts and novices performed equally well when evaluating the incorrect localization by a peer of an axial CT image of the abdomen within the body (novice 100%, Experts 95%; $p = 0.288$), and when evaluating correctness of a peer's assessment of a structure being normal or abnormal ($p = 0.531$). When evaluating the safety of a biopsy approach, both novices and experts agreed when the approach was not safe (100%), however the reasoning behind their decisions differed.

CONCLUSION

Novices perform equally well as experts in recognizing correctness in a peer assessment of correlating medical imaging with the physical body. Novice depth of reasoning may not be as organized and developed as experts. This work provides evidence of different processing occurring in recognition and recall, which may be helpful when designing and assessing learners within a curriculum.

Introduction

Problem solving is a complex cognitive process that is goal centered (Anderson, 1982) and foundational to scientific reasoning. Problem solving is centered around a situation where a problem is identified, and finding the solution to the problem has some value (social, cultural, or intellectual) (Jonassen, 2000). Medicine involves diagnostic problems that require resolution or a potential solution to help the patient, which can comprise all kinds of diagnoses and treatments. The key features in clinical diagnostic reasoning - from data acquisition to diagnosis - are knowledge, context, and experience (Bowen, 2006).

Within radiology, the interpretation of medical images begins with understanding how the image was formed (i.e., the physics underlying image acquisition which creates the unique appearance of a radiographic image – the 2D representation of 3D anatomy) (Manning, 2010). Then the problem solving begins, using the images accompanied by the clinical history and potential prior comparison images. Diagnostic problem solving in radiology begins with the recognition of a radiologic object (something in the image) (Kundel, 2006; Manning, 2010). The next step is to determine the meaning of this object (finding) in the context of the medical image and clinical history (Kundel, 2006; Manning, 2010) – is it a variation of normal anatomy or a possible abnormality. Thus, radiology image interpretation involves skilled decision-making – using cognitive problem-solving skills to make sense out of perceptual information.

Each imaging study in radiology can be viewed as solving a problem (problem-solving). As a radiologist, while interpreting mammograms, I am focused on trying to recognize the appearance of suspicious lesions. At first glance, it is a pattern recognition

test, which is developed over many years of reading mammograms. This is only one process in the detection of cancers. The next step in problem-solving is to integrate those findings in the context of imaging study to determine if it has suspicious enough features to warrant additional imaging and/or tissue sampling. This secondary process, involves greater complexity of thought and drawing upon prior knowledge and experience.

How do radiology medical students come to learn to do this type of problem solving so common in practice? Early in their careers, medical students often learn medical knowledge using a structured curriculum. In the first years of medical school, the students build up knowledge of anatomy and disease processes. When these students enter clinical practice training, the clinical problems are embedded in the patient's story and questions and may involve many organ systems (Bowen, 2006), which is very different from their curricular structure. This transition to clinical practice often finds students awkwardly and slowly recalling basic science knowledge from the classroom and applying it to clinical problem solving. However, over time, the students make new connections between their knowledge and these clinical encounters (Bowen, 2006) and their context and experience are developed and nurtured in the clinics.

Other than this general understanding of the curricular flow, very little is known about how radiologists, and experts in general, develop these problem-solving skills. Problem solving is a skill that involves a combination of recognition and recall of knowledge. This study will use the expert-novice paradigm of research to look closer at novices and experts use of their knowledge of recognition and recall to solve medical imaging problems.

Literature Review

This literature review begins by defining how experts represent their knowledge differently from novices, and then discusses how this knowledge structure may influence components of recognition and recall of knowledge involved in problem solving.

Representation of Knowledge by Novices and Experts

The knowledge differences between novices and experts extend far beyond the aggregate of the knowledge each has acquired for a domain. Understanding the differences in how experts and novices organize their knowledge can bring us one step closer to understanding how they learn, reason, remember, and solve problems. There are many differences in their knowledge; here we will limit our discussion to differences in recognizable patterns, perception and structure of knowledge.

Recognizable patterns (chunks)

Experts have a greater number of recognizable patterns or chunks in memory (Bransford, Brown, & Cocking, 2000; Chi, 2006a; Gobet & Simon, 1996). Each of these 'chunks' has a larger number of parts. During a given exercise, novices and experts recall chunks of memory that apply to the situation. The difference between experts and novices may not be in the number of memory chunks recalled, but rather in the quality of the chunks of memory (Chi, 2006a; Gobet & Simon, 1996). Experts recall higher order chunks of memory first, which also leads them to being faster and more accurate. Experts have more information that is of higher quality and better organized than novices (Chi, 2006a; Young, Van Merrienboer, Durning, & Cate, 2014).

For example, when evaluating a mammogram, novices (radiology residents) will create a list of the findings (e.g., bright area, vague asymmetry) without any relevance to the importance or meaning of the findings. The same mammogram evaluated by an expert radiologist, would stress the suspicious finding (e.g., suspicious mass) and the accompanying salient features (e.g., spiculations) directing the diagnosis and management (medical decision making). The expert packages the information about each lesion, rather than listing features seen on the mammogram.

Perception

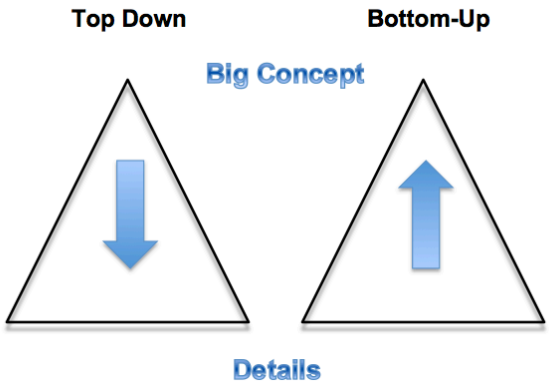
Experts perceive things differently than novices. It is not a matter of just seeing more, but perceiving more (Chi, 2006a). Experts have enhanced sensitivity to critical cues, features and dimensions that novices are not attuned to (Chi, 2006a; Snowden, Davies, & Roling, 2000). In turn, experts are also able to relate those features to a diagnostic explanation (Chi, 2006a; Lesgold, et al., 1988).

Kundel and colleagues (2007) evaluated experienced radiologist and training radiology residents while looking at sets of normal and subtly abnormal mammograms. They recorded the search time required to locate the first cancer and the initial eye scan path. Experts have a global perceptual process that analyzes the visual input of the entire retinal image that is often able to locate the cancer in less than one (1) second of viewing, which is much more efficient than search-to-find strategies of novices.

Structure of Knowledge

The knowledge of experts is organized around overarching principles and higher-level concepts. Experts have knowledge organized at the domain/subject level, which is critical to solving the problem, where novices, will have similar items sorted based on literal surface features (Chi, 2006a), all operating at a list level with little hierarchy to the items. Expert's knowledge is more fully developed and differentiated at both the subordinate and superordinate levels (Chi, 2006a).

Schmidt and colleagues (1993, 1990) examined this organization of knowledge in medical education and proposed the concept of 'knowledge encapsulation'. Knowledge encapsulation incorporates all lower level detailed schemata, concepts, and their interrelations in a combined collection under a smaller number of higher-level schemes with the same descriptive power (Barz & Achimas-Cadariu, 2016; Schmidt & Boshuizen,



1993). Experts with their years of exposure to clinical cases, use short cuts in making diagnoses facilitated by 'encapsulated knowledge'. Novices approach these same problems from the bottom-up, through a process of collecting and interpreting the signs and symptoms. This bottom-up approach of novices will eventually lead to a growth of their knowledge base (Barz & Achimas-Cadariu, 2016; Schmidt & Boshuizen, 1993). Through

repeated exposures and applications of acquired knowledge, novices will begin to change their knowledge structure, and eventually encapsulate knowledge into these meaningful packages that have been developed by experts (Schmidt & Rikers, 2007).

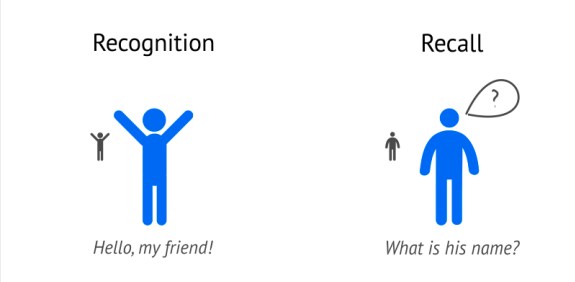
We have discussed the ways that knowledge is organized for novices and experts. The literature overwhelmingly shows that the experts performing significantly better on tasks due to their knowledge organization (Chauvin, 1988; Chi, 2006a; Dunphy & Williamson, 2004; Johnson, 1988; Lesgold, 1983; Myloppoulos & Regehr, 2009; Rikers, Schmidt, & Boshuizen, 2000). However, problem solving is not just about what knowledge we have, but also how we are able to use our knowledge. The literature describes two different ways that we use our knowledge – recall and recognition.

Recall versus Recognition

We've all had the experience of trying to remember the name of a movie or song we've forgotten. We can't recall it no matter how hard we try, but when someone suggests the name, we can clearly recognize the name of the movie or song. One way cognitive scientists have talked about this difference in memory structure is by using the concept of recall and recognition.

We define recall as the ability to retrieve content from long-term memory, and recognition as the ability to identify it based on a familiarity. Both recall and recognition rely on prior exposure to the entity being called upon. For example, if a person is shown an axial CT image of the abdomen, they will likely *recognize* the organs present in that image – stomach, spleen, liver, etc. However, to understand where that slice is located in the body,

the person will have to *recall* prior knowledge relating to where those organs are with respect to the physical exam and also their relationship to the surrounding organs so they can correctly localize the image within the body. The following images describe this situation in a more real-life situation. One may encounter a person on the street and 'recognize' them as a friend; however, it requires 'recall' to generate that person's name.



(Image from webdesign-review.blogspot.com)

Recall

There are two types of recall that are typically investigated in memory testing. Cued recall is recall driven from providing a list to a subject and then having them recall an item off that list with cues to aid the retrieval process (Raaijmakers & Shiffrin, 1992). Free recall is more complex than cued recall (Raaijmakers & Shiffrin, 1992; Anderson & Bower, 1972). Free recall, involves both short term memory and long term memory working together to retrieve information. The initial recount of the item in triggered in short term memory, and the prior knowledge is retrieved from long-term memory working together in a search/retrieval strategy (Raaijmakers & Shiffrin, 1992; Craik & Lockhart, 1972).

Open-ended questions and some multiple-choice questions can be considered recall assessments of knowledge in educational context (Custers, 2008). True-false questions promote recognition, and multiple-choice questions call upon a mixture of recall and

recognition (Custers, 2008; Arzi, Ben-Zvi, & Ganiel, 1986; Anderson & Bower, 1972; Joorabchi & Chawhan, 1975).

In clinical practice, the clinical teams make regular patient rounds of the hospitalized patients. The attending will often ask the resident caring for the patient to present the patient's case to the clinical team. This resident will need to recall in great detail all the pertinent and defining symptoms and signs of clinical presentation and past medical history and relate this to the current medical problem that the patient in being treated for during this hospitalization. The attending then will often ask further probing questions related to the diagnosis, for which this resident will need to recall the specifics related to that disease and relate them to this patient. In this scenario, the resident is drawing upon the current clinical situation of this patient and applying to their past knowledge of disease processes that they have learned from prior clinical experience and study.

Although it is not called recall, much of the research performed in medicine and radiology imaging focuses on recall. For example, Lesgold and colleagues (1988), asked experts and novices to look at chest x-rays and examine the finding, eventually dictating a report. Later they asked them to justify their diagnosis by drawing a line around the finding on the x-ray. The experts were more efficient and made more accurate diagnoses. Similarly, Crowley and colleagues (2003) studied novice and experts ability to make microscopic pathology diagnosis on slides of breast tissue. The experts and novices differed in their search patterns, speed, accuracy and rate of errors. Notice that these studies all focus on recall of knowledge. These studies all demonstrate that experts are better than novices at recall.

Recognition

Recognition is simply defined as the ability to identify something. There are two distinct processes acknowledged in the literature for recognition memory: recollection/remembering and familiarity/knowing (Brandt, Cooper, & Dewhurst, 2005). Recollection is recognition that is accompanied by specific contextual information about a study item when presented with that item. For example, recognizing the location of an item on a study list, or thoughts and images that come to mind when a participant is presented with an item (Brandt, Cooper, & Dewhurst, 2005). Familiarity recognition lacks episodic detail. It arises when someone just knows that an item has been previously presented (Brandt, Cooper, & Dewhurst, 2005).

In radiology image interpretation begins at the level of perception and recognition. The field of image perception is vast and is beyond the scope of this discussion. A lower extremity x-ray from a 20 year-old male with new increasing pain without trauma is presented to the radiologist. The radiologist *recognizes* that there is an ill-defined lytic lesion in the proximal tibia. However the mere recognition is not enough to help the patient's clinician to know what to do for the patient. The radiographic features of the lytic lesion, along with the age of the patient and the location in the tibia will bring about the *recall* of a list of diagnostic possibilities: eosinophilic granuloma, infection Ewings' sarcoma and osteogenic sarcoma (Helms, 1995). The items from this differential diagnosis range from benign to malignant, with a wide range of treatments. The radiologist will need to further *recall* knowledge to help differentiate this lesion, which may require additional imaging tests and/or biopsy to lead to the eventual diagnosis.

There are many debates on how recall and recognition differ and if there they are one, two or many step processes (Kintsch, 1970, Tulving & Thomson, 1973; Haist, Shimamura, & Squire, 1992). However, one thing that the debates have in common is that recognition is a simpler process than recall (Tulving & Thomson, 1973; Tversky, 1973; Kintsch, 1970). Recognition is easier that recall because it bypasses the retrieval process (Anderson & Bower, 1972; Tersky, 1973; Kintsch, 1970).

Research Question

Prior work in medical education (Bransford, et al., 2000; Chi, 2006a, 2006b, 2011; Ericsson, 2004, 2007, 2008; & Lesgold, et al., 1988), including my own (Salkowski & Russ, 2017), indicates experts are better at recall, likely due to their increase in knowledge and how they organize this knowledge. However, much of this prior work has been examined these differences based on recall type tasks. For example, Azevedo and colleagues (2007) investigated the differences in mammographic interpretation of increasing levels of expertise (medical students, radiology residents and radiologists) through the use of a computerized tutor. They observed an increase in the number of radiological observations and findings, and correct diagnosis with the increase in expertise. This investigation reflects a recall of knowledge. Similarly, Hmelo-Silver and Nagarajan (2002) compared the scientific inquiry skill of experts (cancer researchers) and novices (medical students) in their ability to design a Phase 2 clinical trial of a new investigational cancer drug. Both groups had to rely on the recall of past knowledge and critical reasoning to develop an experimental strategy. The groups had similar endpoints, however their reasoning differed considerably likely due to their level of expertise.

If recognition is easier than recall, could novices perform as well as experts in recognition tasks? Peer-based assessments occur on a daily basis in medicine. When seeing a patient, the patient's past clinical diagnoses and treatments are reviewed and correlated with the patient's current complaints and symptoms. In radiology, this occurs with an initial visual assessment of the medical images, which is recognition of the imaging study, anatomy present and any derangement present. Once a finding is noted, this will be explored more deeply, but that first exposure to the images is recognition. If a colleague previously interpreted this imaging, their written report will be used as a peer assessment of my initial recognition. This process is a quick peer assessment if the impressions agree, however if they differ this usually involves exploring the case and the supporting facts in greater detail, which goes beyond recognition and involves recall and advanced critical thinking skills.

Research Question: How do novices and experts evaluate peer-based assessments on medical imaging examinations?

To answer this question I conducted an expert-novice study using the lens of recall versus recognition to evaluate peer-based assessments of medical imaging examinations.

Methods

Participants

Ten radiologist experts who interpreted abdominal imaging as part of their practice (average age 47.4 years with 13.5 average years of experience; 9 males, 1 female) were

invited to participate. Radiologists who interpret abdominal imaging were chosen to exclude non-expert bias from radiologists that were sub-specialized in other fields such as neuroradiology, musculoskeletal, chest or breast imaging. Six of the radiologists were community based and four radiologists were academic based within the same large Midwest University radiology practice group.

Eleven senior (4th year) medical student novices from a large Midwest medical school (average age 27.4 years; 5 males, 6 females) were invited using a snowball invitation method. Students from the author's advanced anatomy course who participated in a pilot study using the simulation tool invited peers to participate. These contacts were sent to an administrative assistant who arranged the session times and dates.

This study was IRB approved. All participants, experts and novices, were compensated with a university logo travel mug.

Study Activities

Since it would be difficult to collect controlled data of radiologists and novices looking at actual images in a clinical setting, we used a simulation tool that was developed for these types of scenario evaluations. The simulation tool allows the user to control a calibrated CT scan of a human torso by moving a handheld probe over the torso and viewing the 'real-time' images on an iPad. The integration of the handheld probe controlling the images allows for direct correlation of the physical exam and the imaging at that location (Figure 1 and 2).

This study then evaluated novices and experts doing tasks with the simulation that are modifications of familiar tasks. The medical images were real, but the ability to

navigate a CT scan of a patient (viewed on the integrated iPad screen) by moving a handheld probe over a simulation torso is in the unfamiliar category. The movement of the probe is similar to using an ultrasound probe, however the images being navigated are CT images. This task was equally unfamiliar to both radiologists and the novices (students). This task allows us to compare the representative knowledge in a 'familiar' clinical context and ask questions to determine similarities and differences between novices and experts.

The participants were each given the same sequence of CT images (Figure 3 and 4). Participants were allowed practice time with the simulation tool, followed by the study session. The same image set was used for all participants to allow for correlation across expertise.

The simulation session to evaluate peer knowledge was divided into three segments:

Segment 1

The first segment consisted of the participant being shown a CT image on the iPad screen and the placement of the probe on the stimulation torso. They were told that the probe location represents the location where their peer (medical student for novice, and radiologist for expert) placed it. They were asked to assess the correctness of the placement to the CT image presented on the iPad (higher, lower, or at the correct site). In this task, the participant must recognize the anatomy that is present in the image, and secondly they must recognize where the location of the probe was put on the simulation torso. If they felt it was not correct, then they were instructed to 'demonstrate to their peer' where the probe should be located (Figure 3). The target location – the correct

location for each image - was defined within the simulation by the image preset for each case. The location of where participants placed the probe was recorded and compared to the target location. Once localized, they logged in that position of the probe by tapping a button on the iPad. Within the simulation, the movements of the probe and initial/final localizations were identified when the participant locked these movements on the iPad. At that point they were presented with an image of where they placed the probe, side by side with the target image. Each participant did two events in this segment.

Segment 2

The second segment was assessing if the structure within the CT image that they were shown on the iPad was normal or abnormal. In each case, they were given the assessment, normal or abnormal, that their peer had given that anatomical structure (Figure 4). Their task was to agree or disagree. In this task, the participant must recognize the organ that they are being shown and be able to recognize its normality.

Segment 3

The third segment consisted of evaluating a biopsy approach for the right kidney. The participant was to act as the instructor overseeing the potential biopsy procedure and assess if the anterior abdominal biopsy approach presented should proceed (Figure 5). In this task, the participant must recognize the anatomy that they are being shown in the image and the approach of the presumed biopsy. If they answered to stop the biopsy, they were asked if there was another solution available and why it was favored over the approach presented. Likewise, if they answered to complete the biopsy, they would be asked details about the biopsy procedure. The same axial CT image set was used for all participants to allow for correlation across expertise.

All participant decisions were made at a peer level to facilitate cognitive congruence and a safe environment to provide judgment (Ten Cate & Durning, 2007; Helfer, 1972). It also simulated the peer-to-peer interaction, professional competence, and referral process that occur in daily medical practice (Norcini, 2003). In all situations, it was said that a prior peer made this evaluation, rather than using the proctor as the 'simulated' peer to avoid the in-person confrontation. The intention was to create a same level of evaluation, rather than forcing the evaluation of a higher authority. Peers are able conduct reliable and valid evaluations of their peers (Arnold, Willoughby, Calkins, & Gammon, 1981; Helfer, 1972), and it may provide insight into their own level of knowledge (Linn, Arostegui, & Zeppa, 1976).

Data collected

Simulation session data

During segment 1, the location of the target and the 'corrected' location by the participant were recorded. This allowed assessment of how close the participant placed the probe to the target when correcting their peer's probe location. The data from the simulations was recorded through the ADAGE computer program specifically developed for this simulation.

Think Aloud data

While the participants were engaged in the three (3) tasks they were asked to verbalize their thoughts. Think aloud is a method of gathering information that provides

rich verbal data about an individual's reasoning during a problem task. It provides the in-the-moment information about what is being concentrated on and how information is structured during a problem-solving task (Chi, 2006a; Ericsson & Simon, 1980; Fonteyn, Kuipers & Grobe, 1993). This method provides valid data on the underlying thought processes that are occurring during the activity (Ericsson, 2006). Ericsson and Simon (1980) realized that when subjects are working under a heavy cognitive load, they tend to stop verbalizing or provide less complete verbalizations. In cases when participants neglect to verbalize, they are asked questions related to that event right after completing the task. The think alouds were audio-recorded and later transcribed.

Data Analysis

Qualitative data

The transcripts and data collection sheet used during the simulations were analyzed based on the simulation segment. In segment 1, record was made of the 'peer-placed' localization, and then the evaluation made by the participant. In segment 2, the participant's evaluation of their peer's abnormal/normal decision was recorded. In segment 3, due to the greater complexity of the task the transcripts were coded for four main themes (biopsy safety, reasons, alternate approach offered, and other pertinent information). These codes were derived from the questions asked during the biopsy evaluation task. The results were tallied by level of expertise and grouped according to similar codes (Tables 5 and 6).

Statistical analyses were performed with the on-line Chi-Square Calculator by Social Science Statistics (2017). The association between expertise (expert and novice) and

assessment correctness, expertise and peer agreement, and expertise and biopsy safety were analyzed. Significance was assessed at $p < 0.05$.

Quantitative data

During segment 1 of the simulation, participants evaluated the location of where a peer placed the probe for the presented target image. In each case the true location (target location) was known to the proctor, and recorded on the log sheet. The location of where the peer placed the probe was also recorded on the log sheet as being higher or lower than the target location (truth). The participant's initial evaluation of where their peer placed the probe for the presented CT was recorded (higher, lower, or on target). Finally, the probe location of where the participant demonstrated to their peer where the probe should be located was recorded within the simulation. The correctness of these responses was compared with the target location. The correctness of the evaluations (i.e., distance and direction between probe and true location) was compared with independent t-tests calculated with SPSS, using guidance from Laerd Statistics (2015). Significance was assessed at $p < 0.05$.

Results

Recognition

The different tasks in the interview allowed us to identify a variety of ways in which experts and novices compare in evaluating peer assessments. Below we look at the

correctness of the location, determine if the structure was normal or abnormal, quickness in making a judgment and evaluate if the biopsy approach was safe.

Both experts and novices were able to make judgments about their peers' localization – that is, they were able to recognize errors (or not) in the work of their peers. For localization, novices provided correct evaluations (i.e., assessed the probe too high for target when it was too high, and assessed probe placement too low when it was too low for the provided CT image) 100% of the time and experts 95% (not statistically different; χ^2= 1.1268; p=0.288) (Table 1).

Novices and experts were equally proficient in evaluating if a structure on the CT was normal or abnormal as evaluated by their peer (i.e., if their peer evaluated the kidney as normal they would agree or disagree, if they evaluated the same kidney to be normal or abnormal). The experts agreed with the peer assessment in 75% (15/20 cases) and novices 82.6% (19/23 cases) (not statistically significant; χ^2=3.5643; p=0.168) (Table 4). In all cases, the organ/structure they were asked to evaluate was normal, however, their peer may have evaluated it as normal or abnormal. The participants, had to evaluate the normality of the structure, and also determine if their peer made the correct assessment.

When novices and experts were making the evaluation about their peer's localization for the given CT image, their responses were equally as fast, 60% of experts responded quickly and 68% of novices (not significant; χ^2= 0.3055; p=0.580) (Table 3). The quickness of evaluation was coded from transcripts. A quick response was coded if the participant responded with an answer right after the question, and delayed response was

coded, if they paused, 'umm'd', asked questioned prior to giving a response, or gave explanation about reason prior to answer (no formal time data was collected).

Additionally, the transcripts were coded for explanations of how they came to the decision of evaluating the correctness of the peer's localization. Three categories emerged: regional anatomy references, past use of the simulation tool, and prior clinical experience. Both novices and experts used comparative regional anatomy the most to explain their decision (Table 3). That is to say, that the decisions in evaluation were based on their recognition of the surrounding anatomy.

Despite no difference in quickness of evaluation or their use of reasoning, there was a difference in how and if they provided a reason. Experts were more likely not to provide any reasoning for the evaluation (45%), or they needed prompting to provide their reasoning (40%). Only 15% of expert evaluations freely offered reasoning. Compared to novices this was a significant difference (χ^2= 6.1409; p=0.046). Novices freely provided a reason for their evaluation (41%), needing prompting in 45%, and 14% did not provide a reason. The fact that the experts did not provide reasons or needed to be prompted for a reason for the evaluation in localization segment may suggest to an expert that the answer is already implied and no further reasoning is needed. This may be a reflection of their level of experience.

Novices had two (2) cases when they were not able to call the structure abnormal or normal (Table 4). The novices reasoning for not being able to make the judgment were, 'I've never really seen an abnormal stomach', and 'No, I just see gas in the stomach, which is normal. I mean, maybe -- I don't know. I don't know what could be abnormal.' In the end,

the novice would not commit to the assessment of normal or abnormal. The experts were able to make a yes or no type judgment in all cases.

Lastly, when novices and experts were asked to determine if the approach to doing a renal biopsy (from the anterior abdomen) was a safe approach, both novices and experts said that it was not safe (100%) The correct answer was that the biopsy approach was not safe in normal conditions. Three of the experts did volunteer that the anterior approach could be made if there were no other approaches. Flin and colleagues (2009) noted that medical students are able to identify medical errors, but are often unsure what to do if a colleague made an error.

These examples represent the process of recognition. Both the novices and experts were able to evaluate the correctness of the location, determine if the structure was normal or abnormal, and evaluate if the biopsy approach was safe. These questions are a form of a yes/no or true/false, which more closely represent the assessment of recognition. Both novices and experts were familiar with the anatomy enough to impose a judgment with reasonable confidence, without having to call upon deeper knowledge to further explain the reasoning or go into more detail.

Recall

We have shown that novices and experts can equally well recognize the correctness of a peer's localization. However, going beyond the yes/no type recognition to show the peer the correct position requires retrieving more knowledge and applying it to a situation. That is, it requires recall. How do experts and novices compare in recall based on their recognitions?

When demonstrating to their 'peer' the correct location for the probe placement on the body corresponding to the target image, experts and novices performed similarly in this task (Table 1). However, experts tend to be closer to the target at final placement (Table 2) regardless if the probe placement is higher (towards the head) or lower (towards the feet) than the target. This probe place was significantly better (closer to the target) for experts ($p=0.041$) when the localization was made lower on the body than the location of the target. Similar better results for experts were seen when the probe was localized above the target for experts, however this was not significant.

These findings suggest that the experts may be able to call upon spatial knowledge to more closely localize the image when 'demonstrating' to a peer. Performing this task goes beyond recognition. When localizing the position of the CT image within the simulation torso, there is a component of the mechanical skill. Because this mechanical skill of using CT images in this manner is unfamiliar to both novices and experts, they have the same disadvantage. The ability to localize a structure spatially relies on the ability to recall the adjacent anatomical structures to decide if the target image and its composition of structures are above or below the surrounding anatomy (in the axial plane). This could be analogous to the master chess player's ability to recall the sequence of moves played on a visual display of a chessboard (Chi, 2006a; DeGroot, 1965, Bransford, et al., 2000). The expert is drawing on their 3D spatial knowledge and their ability to perceive critical cues that the novices were not attuned to. This superior recall of the experts is the result of being able to 'chunk' their spatial knowledge in meaningful patterns of information; likely from looking at thousands of CT scans in the clinical setting.

In addition, when novices and experts were asked to evaluate a medical procedure (biopsy task in segment 3), they both recognized that the procedure was not safe. Beyond this yes/no question, there were differences in how they solved the 'better' biopsy approach. The novice knowledge structure demonstrated a more bottom up approach and the expert's approach was a top down approach (Table 5). The experts also perceived critical cues that novices did not notice or see as relevant.

Here is an example of a novice (representative of novices in this study) who was concerned about the details about the biopsy approach and why this is not the best approach:

'So I'd say look at where you'd be going in right where you start and where your target is. And look what structures you might pass through, and that's probably the liver there. Well, that's the liver there, and you -- I mean, in all likelihood go straight through the lower edge of the liver……. Uh, I mean, you wouldn't want to hit the liver and, I mean, damage it, but at the same time if you're going for a biopsy, then you're just going to be biopsying the incorrect thing as well, so you're not going to get a good biopsy either. But mostly to prevent the patient from undergoing harm that you could prevent by missing the liver.'

And for the alternate biopsy approach offered by this novice:

'So I would -- I think that you'd want to go, like, in with a posterior approach, because if you're looking at this scan, there's less structures to pass through in

the body. I mean, you're going to have to damage something, obviously. But it I think a posterior approach, probably a little bit lateral and trying to minimize the amount of structure to go through.'

In comparison, this expert (representative of experts in this study) is evaluating the situation with a more global approach and reasoning through the procedure and possible complications:

'You can biopsy through the liver. You certainly can. But if you could, you would try to get the liver out of the way or angle your approach so that you would go sub-hepatic. And you could get the patient to [INTAKE OF BREATH] take a breath in, blow it out, and hopefully that liver will go up high enough. So I would want to see the patient, you know, breathing. And in the appropriate breath stage of respiration, then I would actually try and go in there.'

And for the alternate biopsy approach offered by the expert:

'I would go from behind actually because, you know, if you can get the patient to take a breath in and just get a posterior. I mean, there is much less damage that you can do posteriorly if you go subcostal. So I would actually, you know, tell the patient go prone.'

This contrasting example for the novice and expert approach demonstrates that the novice was only concerned about the liver, and was not able to provide any details for a

procedure beyond the issue of the liver. He/she became stuck in that one issue or detail. His/her lack of experience affected his/her misinterpretation of cues. Because the lack of knowledge for performing this type of biopsy, he/she were also inefficient in their approach to the problem. This type of reasoning was typical of novices. In contrast, the expert acknowledged the liver, but quickly provided options to deal with the liver if that had to be the biopsy approach.

Similar findings are noted when the novice and expert offer an alternate approach to the biopsy. They understood that the approach was incorrect, but could not formulate a concise plan for a biopsy approach other than the need to go posterior to avoid some structures. In contrast, the expert offers a reason and solution to the approach even with commenting on how to position the patient. These types of answers were typical for expert and novice alternative approaches.

This example demonstrates that the recalled information by the expert is more organized around the anatomy and procedural goal than for novices. Coming at the situation to solve the problem, the expert draws upon past experience including the details of breath holds, approaching the biopsy from a subcostal location and positioning of the patient (prone). The thought structure by the novice reflects that their knowledge is not structured around past experience, likely because they lack experience in performing or even observing such a biopsy. Instead, they must rely on their prior anatomical knowledge to make these judgments.

Looking across all the novices and experts, the novices had a list of six reasons and experts had four (Table 5) reasons for not doing an anterior approach for the renal biopsy. The reasons provided by novices could be classified as ‘details’ (for example, would go

through the liver, lots of bleeding, injury to intestine); versus the reasons provided by the experts were more 'big concept' (for example, avoid liver, injury to intestine or vascular pedicle of the kidney, and don't want to go from peritoneum to retroperitoneum). This example demonstrates that the novice's knowledge is organized around details, unlike the experts who have organized their knowledge about the procedural anatomical approach into larger concepts – 'larger chunks'.

Discussion

Despite the fact that experts and novices normally show large differences in most studies, novices and experts performed equally well in this study. Most research studies on novices and experts focus on the knowledge content and the differences in the organization of the knowledge between these two groups. These investigations focus on recall of information which overall is a more difficult operation. It is no surprise that the experts would out perform the novices in these tasks or measurements. It should not be disregarded that easier tasks such as recognition are less important.

Novices in our study preformed equally well compared to experts on the recognition of an improper localization by a peer, the recognition of a normal/abnormal anatomical structure and the recognition of the safety of a biopsy approach. They are likely able to do so because of their prior exposure to anatomy, medical imaging and physical examination skills. Although they don't have the depth of knowledge and experience of the experts, their current knowledge allows them to recognize important principles and concepts. The fact that the novices and experts perform similar in recognition tasks is surprising with the vast prior research showing that experts out perform novices, however, there may be

something unique about recognition that does not require the vast amount of knowledge or the complex organization of knowledge of expertise.

Our findings here are consistent with those found elsewhere in the literature. Evans and colleagues (2011) found that there was no significant difference between novices and experts (radiologists and cytologists) for recognition of scenes or objects. Vogt and Magnussen (2005) found that artists were equally as good as novices in visual recognition memory for sporting events or abstract paintings. Although these studies differ in the methods, in that participants were shown images and later asked to recognize if they had seen it, these collective studies do demonstrate that novices in certain recognition tasks can perform equally well as experts. Studies that have demonstrated differences, found that experts found the lesions on radiographs faster than novices (Nodine, et al., 2002; Christensen, et al., 1981). In our study the novices and experts had unlimited amount of time to make their judgments.

This study accepts the null hypothesis that novices can perform equally well to experts in recognition. In contrast, prior evaluation of novices and experts in simulation exercises discussed in chapter 3 ('Expert and novice differences in correlating anatomy and imaging') did identify differences in performance on recall tasks. Experts were faster at making decisions on medical images. Experts also used higher order knowledge when localizing images in the body such as relying on smaller or organ substructures, and more accurate terminology. Experts notice patterns on medical imaging not common to novices. These prior results demonstrated that experts outperform novices, yet, within the context of medical imaging exercises involving recognition, novices were equally proficient to experts.

Conclusion

In the medical literature, a novice is described as '*someone that is completely new to an area and who does not possess a significant amount of pre-existing knowledge or skills*' (Dunphy & Williamson, 2004). In medical school, medical students are novices and often treated as if they are deficit of knowledge. They don't have the skill set to be caring for patients. Because they don't know what an expert knows, then they are often casted as not knowing.

However, students (novices) performed equally well to experts in this evaluation of recognition applied to a medical imaging simulation. This raises questions about whether we should treat novices as blank slates as we often do in medical school when they enter clinical practice.

Recognition is easier than recall because it bypasses the retrieval process. Experts are able to excel in the recall due to their developed skills in recognizable patterns, perception and structure of knowledge. However, the fact that novices can perform equal to experts in recognition is a foundation upon which learning can be built. Recognition can lead to recall provided there is underlying structure and knowledge. Exposure to more situations of recognition can increase or at least help promote the linkages to underlying knowledge and thus foster recall.

Medical students excel in recognition of medical imaging findings in a simulated clinical setting. Rather than focus on their inability to formulate a complete diagnostic evaluation, it may be more helpful to approach learning from development of their prior knowledge (recognition). Some approaches could be assessing student's prior knowledge at the beginning of courses to determine the amount of content that needs to be

reintroduced, versus continued development of the material. Because the students are good at evaluating peer based answers, this may be reason to introduce peer-based education as part of a curriculum design. Problem based learning or case-based problems, may also help to develop knowledge from a level of recognition to structured knowledge that can be easily retrieved.

The fact that novices in this simulation exercise perform equally well compared to experts in recognition has implications for how we can approach clinical teaching especially in radiology with medical students. When discussing clinical imaging cases in a conference or in the reading rooms, we can start by discussing what they do know (i.e., the anatomy), and then proceed with showing the derangements of that anatomy and finally discussing what that means clinically. Thus building upon what they know and strengthening some other knowledge that they know (i.e. pathology curriculum), but not previously seen in this context. This also has future implications for novice-expert research. Studies could be developed that explore other scenarios of imaging recognition that involves true abnormal cases compared to normal cases. Also, since they have prior knowledge, further exploration of depth of recall for related content. This could also be expanded to radiology residents, who would be at intermediate levels of expertise.

Figures

Figure 1: Components of simulation tool. A, iPad image viewer and control panel; B, handheld probe that moves image set in integration with the torso; C, torso model for physical integration with the image set and handheld probe; and D, mounted box housing all the controls.

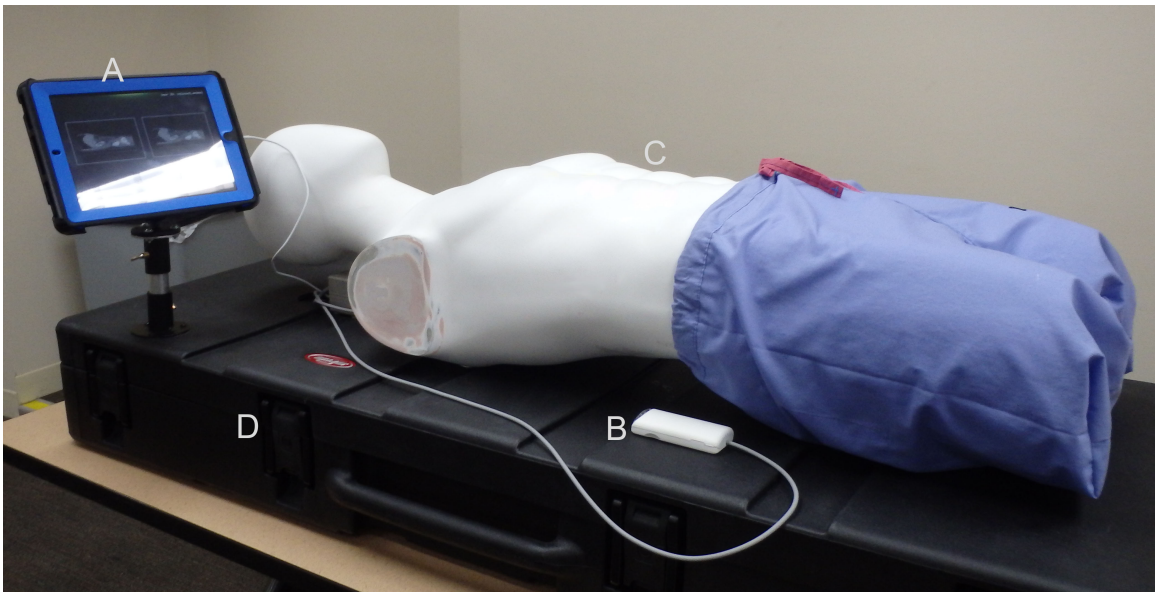


Figure 2: Interaction of the participant with the simulation tool localizing an image within the physical torso.

Figure 3: Image presets (KE2 and KE1) for the participants to assess the correct location of the peer-placed probe position (Segment 1).

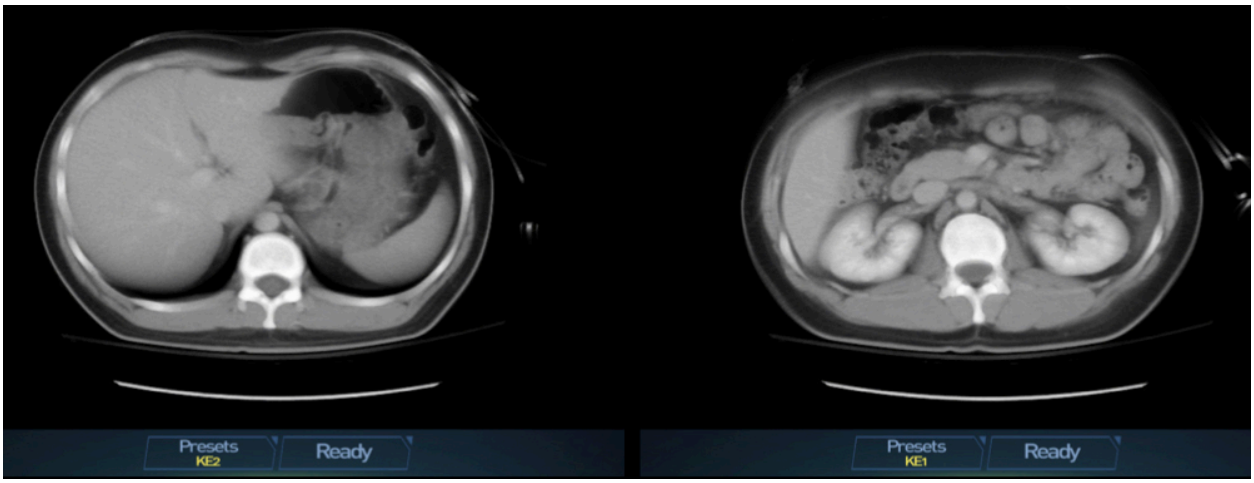


Figure 4: Image presets (KE2-2, assessment of stomach; KE2-4, assessment of right kidney) for the participants to assess normality or abnormality of an anatomical structure assessment of a peer (Segment 2).

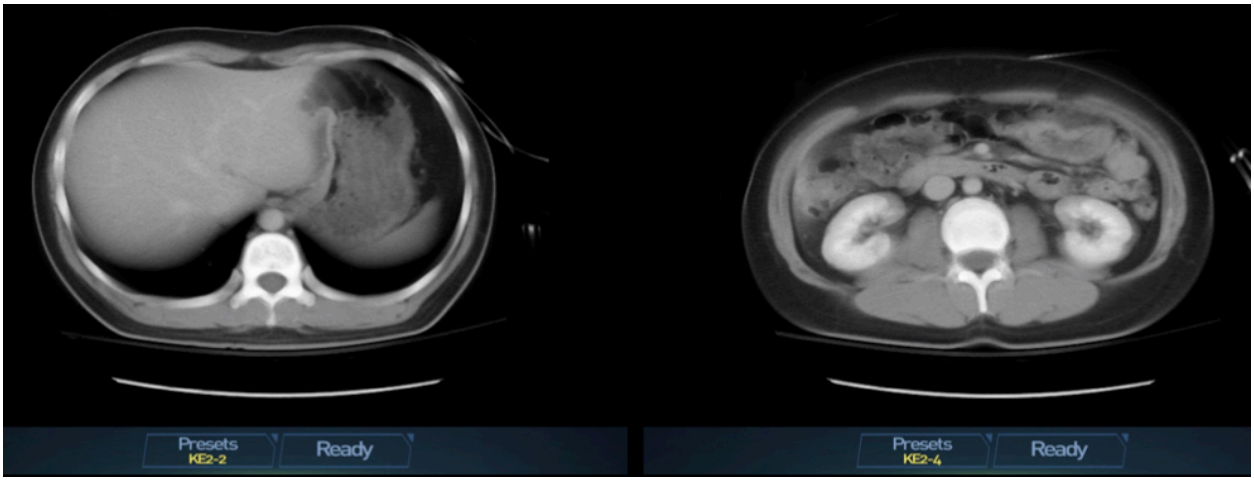
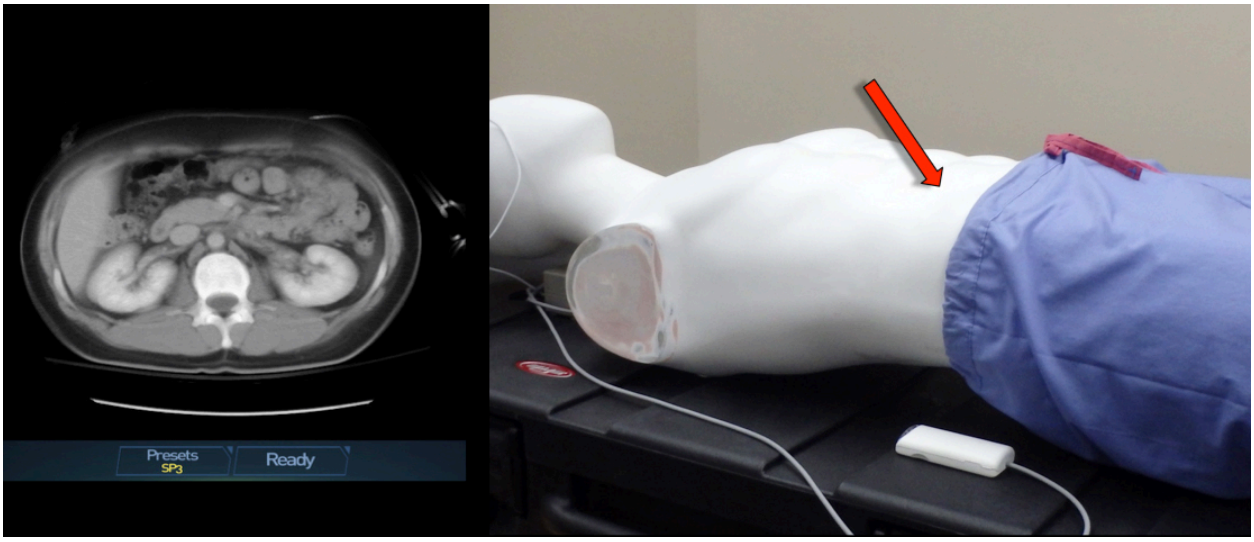


Figure 5: Image preset (left sided image) for right renal biopsy and physical location used to demonstrate on simulation torso (right sided image). The red arrow demonstrates the position and angle of the biopsy to gain access to the right kidney (Segment 3).



Tables

Table 1: Novice and experts assessments of correctness of probe location and then their assessment of where the probe should be placed on the simulation torso for the presented target image (Segment 1).

	Initial assessment compared to peer		Where they placed probe to show 'peer'		
Novice					
Probe positioned low	Assessed correct (low)	13/13 (100%)	Lower than target	6/12 (50%)	**One data point lost
			Higher than target	6/12 (50%)	
	Assessed incorrect (high)	0 (0%)			
Probe positioned high	Assessed correct (high)	9/9 (100%)	Lower than target	6/9 (66.7%)	
			Higher than target	3/9 (33.3%)	
	Assessed incorrect (low)	0 (0%)			
Expert					
Probe positioned low	Assessed correct (low)	11/11 (100%)	Lower than target	5/10 (50%)	**One data point lost
			Higher than target	5/10 (50%)	
	Assessed incorrect (high)	0 (0%)			
Probe positioned high	Assessed correct (high)	8/9 (88.9%)	Lower than target	5/7 (71.4%)	**One data point lost
			Higher than target	2/7 (28.6%)	
	Assessed incorrect (low)	1/9 (11.1%)	Lower than target	0	
			Higher than target	1/1 (100%)	

Table 2: Independent t-Test results for novice and expert assessments of correctness of probe location when instructing peer of the correct location (Segment 1). (The -/+ indicate directionality, with larger absolute value farther from 0)

						95% CI of the Difference	
	Number of localizations too high	Mean distance	t	Sig (2-tailed)	Mean Difference	Lower	Upper
Novice	12	-19.61	1.134	0.272	6.56	-5.59	18.7
Expert	8	-13.05					
	Number of localizations too low	Mean distance	t	Sig (2-tailed)	Mean Difference	Lower	Upper
Novice	9	+17.8	-2.214	0.041	-9.38	-18.31	-0.44
Expert	10	+8.43					

Note: One novice data point lost as task not done correctly, two expert data points lost due to technical Internet issues.

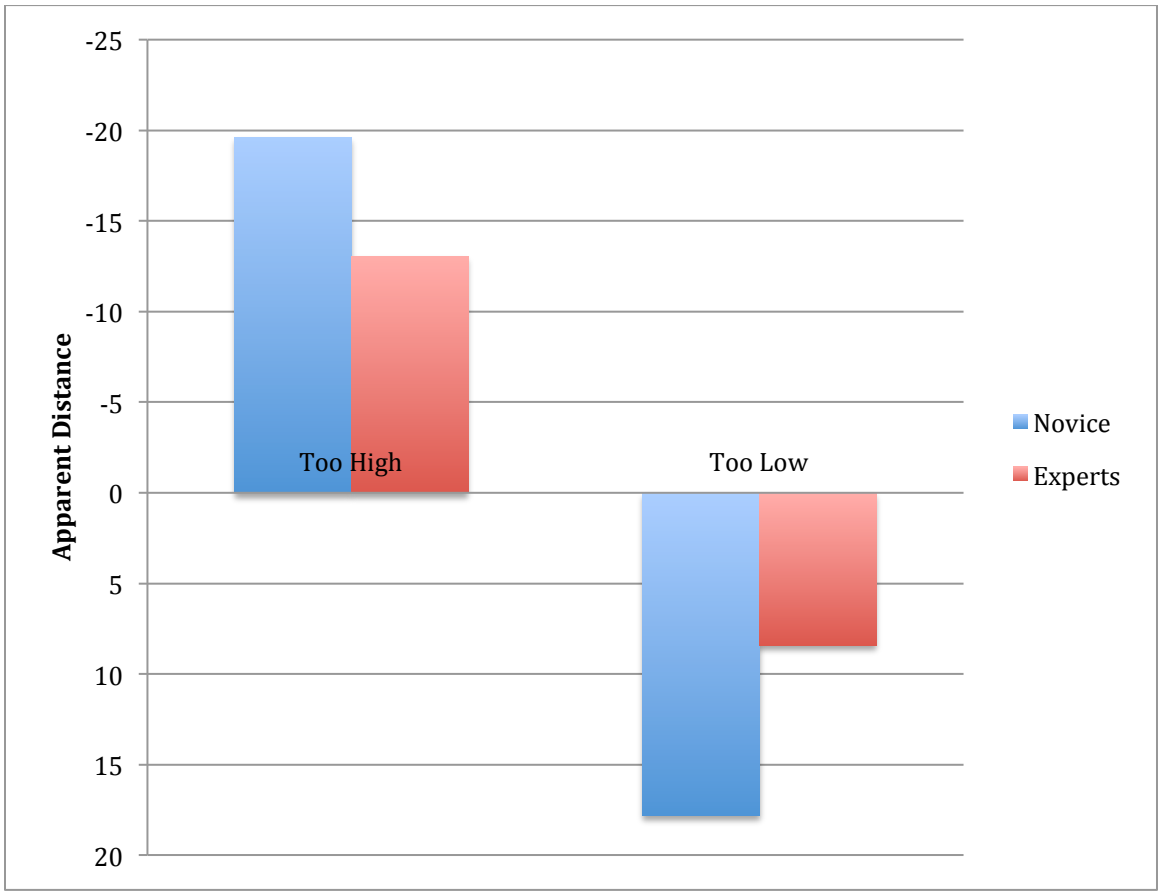


Table 3: Novice and expert responses to evaluation of to the correctness of localization by a peer (Each participant given two evaluations).

Expertise	Quickness in making evaluation of correctness		Reasoning for evaluation of correctness			Approach used for explaining reasoning behind evaluation		
	Fast	Hesitate / slow	No reason provided	Provided without prompts	Provided with prompting	Regional comparative anatomy	Experience with simulation tool	Prior clinical experience
Expert	12 (60%)	8 (40%)	9 (45%)	3 (15%)	8 (40%)	11	1	1
Novice	15 (68%)	7 (32%)	3 (14%)	9 (41%)	10 (45%)	19	2	2

Table 4: Novice and experts agreement/ disagreement with peer an assessment of normal or abnormal (Segment 2). Gray shaded cells represent truth.

		Agree	Disagree	Don't know/can't determine
Novices	Peer labeled abnormal *(truth is normal)*	2	7	2
	Peer labeled normal *(truth is normal)*	12	0	0
Experts	Peer labeled abnormal *(truth is normal)*	2	8	0
	Peer labeled normal *(truth is normal)*	7	3	0

Table 5: Novice assessment of performing a right kidney biopsy from anterior abdomen and recommendations for a different approach (Segment 3).

	Anterior renal biopsy approach not safe	Reasoning for why not to do biopsy	# of replies	Approach recommended	# of replies	Reason for new approach	# of replies
Novice	11/11 (100%)	Would go through the liver	9	Posterior	10	Minimize structures that are traversed	5
		Lots of bleeding with hitting the liver	5			Avoid liver	4
		Injury to intestine	3	Posterior-inferior	2	Only traverse musculature	2
		Injure arterial supply of kidney	1			Avoid vessels	2
		Bleeding	1	Posterior-lateral	1	No reason offered	2
		Enter the peritoneal cavity	1			Kidneys are retroperitoneal	1
						Avoid lungs	1

Table 6: Expert assessment of performing a right kidney biopsy from anterior abdomen and recommendations for a different approach (Segment 3).

	Anterior renal biopsy approach not safe	Can go through liver if no other access	Reasoning for why not to do biopsy	# of replies	Approach recommended	# of replies	Reason for new approach	# of replies
Expert	10/10 (100%)	3/10 (33.3%)	Avoid liver	6	Posterior	7	Minimize structures that are traversed	5
			Injury to intestine	6	Anterior - inferior	2	Avoid liver	4
			Injure vascular pedicle of kidney	3	Posterior-lateral/ posterior oblique	3	Stay in retro-peritoneum	2
			Don't want to go from peritoneum into retro-peritoneum	1	Lateral	1	Avoid intestine	2

Chapter 5

Conclusion and Future Directions

This body of research has demonstrated that simulation techniques can be used effectively to evaluate the differences and similarities between novices and experts. In addition, simulation and specifically the simulation tool developed within this thesis can be used to support the correlation between the physical examination and cross sectional medical imaging.

Novices were able to make correlations between the physical examination and the cross sectional imaging; however this was not as developed as the experts. Novices did have equivalent skills at recognition of normal anatomy or the ability to evaluation the safety of a biopsy approach. Understanding these foundational skills of novices may be the basis on which future curriculum can be developed for making the integration of medical imaging and the physical exam, and likely other challenging areas within the curriculum.

More broadly, this study has demonstrated that simulation can be effectively used as a tool to explore student's knowledge and also apply new educational concepts. Simulation offers a method to identify what a student knows about a concept, and then add newer concepts and principles upon. Simulation events can help to build expertise.

Within education, learning is aimed toward expertise. This study has demonstrated through the novice-expert paradigm that this model can be applied to other teaching methods. Understanding the knowledge base that our learners possess, allows us to build upon that through learning activities. Our learners do not come without prior knowledge and this study demonstrated that we could find ways to capitalize.

Limitations of the Study

The experts within this study were all radiologists, and they have developed skills to effectively correlate the anatomy of the physical examination with the medical imaging. In most areas they excel in comparison to the novices. A missing variable in this study is that we still do not understand how they developed this skill to make this correlation.

Using the large extremes of novice and expert for this study did provide useful data and insights. However, it does neglect the steps that occur between these extremes. Without knowing what or when skill acquisition occurs between novice and expert, makes it difficult to predict how or when to intervene with educational enhancements. Gaining additional steps, radiology residents in the early and later phases of training, between the novice and expert may provide useful information. For future program development of this simulation for teaching the correlation of the physical examination with medical imaging, this sequential development of expertise may be essential.

In the long term, it is uncertain if the ability to make the correlation between the physical exam and medical imaging will have a direct impact on patient care, or the utilization of clinical medical imaging studies. One would anticipate that a better understanding of this correlation would lead to improved patient care, however this may be difficult to validate with the complexity of medical education and educational experiences.

With the changes in healthcare practice, it is important to be aware of policy and procedure changes such that educational resources are implemented at the leading edge. Being at this leading edge serves to provide the biggest impact.

Lessons Learned

Before beginning my thesis, I knew I wanted to do a study with simulation. My internship at GLS during my Masters raised my interest in educational games. Simulation is a great adaptation of educational games, and is something that I can use in my clinical teaching. I am also very passionate about competency assessment in my clinical practice. Thus combing the elements of simulation with more in-depth understanding of the expert-novice paradigm seemed like a great fit. Although this study did not explore competency directly, the ability to investigate the novice-expert framework in this study provided me with insight and foundation for using this framework in developing programs for medical students and radiology residents in the future.

In prior research, I have attempted to use mixed methods, however compared to this study; it was more of a quasi-mixed methods approach. This project really enhanced the understanding of how to use mixed methods in research. I learned early in the development of the project and the clinical interview that the quantitative and qualitative components needed to support each other and not just be add-ons. The analysis of each component needed to be handled separately, but also with keeping the other data analysis in mind. It would have been easier to report the qualitative and quantitative data separately, however that would not have fostered the findings noted in this thesis. It was at times difficult to reason using both analytical methods at once. The qualitative results, provided a deeper meaning to the pure data, and yet the quantitative data added depth to the dialogue. Since doing such a large project I can see this being my method of choice for future educational research projects.

This study produced a lot of data – qualitative and quantitative. At first the table upon table of data seemed insurmountable. Trying to find themes in the data was difficult. Being able to talk about the data and early findings with Rosemary would produce gradated ideas that would eventually develop into themes. This was frustrating and exciting. The iterative process developed a deeper appreciation for research in education. It goes far beyond the numbers, but being able to support the numbers with ideas and theories has been very gratifying.

Additional Findings for Future Inquiry

During this investigation, several unique observations were made but not explored fully for this thesis. This presented at different times during the investigation and had both qualitative and quantitative components.

The first was the issue with scrolling through the image sets by the participants. After they made their initial guess for the location of the image within the body by locking in the location of the probe, they were allowed to move the probe to align it to the target image. They moved the probe over the simulation torso while looking at the iPad screen of the actively changing image with their probe movements. When these probe movements (which I refer to as scrolling) were graphed, there appears to be characteristic movements. The comparison of these scrolling activities needs further evaluation to assess for novice-expert differences, and/or imaging plane.

Additionally, there were not any instructions provided to the participants to touch/palpate the simulation torso during the sessions. The sessions were not videotaped, but only audiotaped and written data notes. However, within the audiotapes and written

data logs, reference to palpation was noted by some of the participants as a way to explain their reasoning. This would be an interesting area of exploration, from a novice-expert approach and in reference to what anatomy was present on the target image.

Lastly, data was collected about a fictitious procedural task. During this task, while only viewing a target image, the participants were asked to identify the anatomical structures that they would traverse while placing a long needle from the area marked on the simulation torso to the area of concern on the target image. This task produced data that included lists of organs to why it was okay to traverse certain organs/structures. This data can be explored for these themes and how they relate to the level of expertise.

Extensions of this Study

This study was centered on exploring the differences of novices and experts using a normal cross sectional imaging study within a simulation tool. There are many directions that can yet to be explored within this paradigm and with some future modifications.

First, it is uncertain if this skill of correlation of the physical examination with the cross sectional imaging is unique to radiologists. To explore this further, this simulation exercise could be evaluated with surgeons and internists, who approach the field of medicine from very different perspectives compared to radiologists, and each other. Surgeons rely on the physical examination and medical imaging studies to guide a surgical procedure, however their view of the anatomy, in my clinical experience of working with surgeons, is different from the radiologist. Internists, in contrast, spend a great deal of time interacting with patients by gathering medical history and performing a physical

examination, and less time reviewing medical imaging. Thus, it would be interesting to view other medical experts within this simulation setting to determine any differences and similarities that may be unique to their expertise.

Second, the fourth year medical students in this study participated in medical school with the traditional curriculum. In the last year, the medical school has undergone a transformation to an integrated curriculum – problem based learning. Problem based learning (PBL) is a problem-centered curriculum where the content is imparted through problems often in small group settings. In this format, basic science information is integrated into the context of a clinical problem or disease. The students are actively involved in their own education and learning, and the process encourages students to become self-directed learners. The integrated basic science curriculum in many medical schools has been compressed into 18 months, thus allowing for transition to the clinics earlier and for approximately 30 months compared to the traditional learning that had only 24 months.

Nandi and colleagues (2000) conclude that the PBL students place more emphasis on meaning than memorization and are more accustomed to using journals and on-line databases over conventional students. They found that the PBL students are more confident in information seeking and use a more in-depth approach to learning. It appears that the PBL students demonstrate better interpersonal skills, psychosocial knowledge and attitudes towards patients. The conventional students do out perform PBL in basic science examinations, and both groups perform equally well on clinical examinations. Overall the PBL students have more motivation. They felt engaged, whereas the conventional students found learning to be "non-relevant, passive and boring". Nandi and colleagues (2000)

claim that a program that combines conventional and PBL may provide the most effective training for medical students, however they do not substantiate this claim.

Epstein (2004) presents the argument of traditional education being increasingly replaced by PBL. He notes that there has been an explosion of PBL and 'clinical competence' (so called just in time learning) at the same time that there is an explosion of information. There is an ongoing debate over the optimal balance between factual (content-based) and practical (training or performance based) components of professional development. Epstein (2004) contends that there has been a switch of the educational philosophy to non-didactic methods as a strategy to handle to the expansive information environment. He questions if this trend to context dependent learning will prove effective. Further he states 'what is the hard evidence indicating that the original educational system was broken and that the new system is likely to fix it?' Epstein also questions if there are long-term hidden costs payable for the short-term benefits afforded by the PBL teaching philosophy.

To expand upon whether this new curriculum has an effect on the correlation of the physical examination and medical imaging, there is a future opportunity to repeat this research on fourth year medical students. Since the medical school recently made this conversion, this research study can first be conducted in three years. This is an opportunity to compare students from two curriculum systems in one clinical application.

A possible limitation to this comparison study between curriculum systems, is that this study is done on top students and thus the small differences in achievements may be more of a reflection of the students themselves than the educational process (Colliver, 2000). Medical students are considered the top students in the class of traditional learners,

thus comparing students that have mastered traditional education to PBL may be an unreasonable expectation (Albanese, 2000). These students are already at the top of the pyramid and expecting higher gains from a new teaching methodology may be limited by the fact that they cannot go much higher. Potentially a difference could be seen if the students compared had the respective learning environments for their entire educational process leading up to medical school.

Third, the simulation tool currently contains a normal male cross sectional imaging study. There is the potential to add clinical pathological imaging cases with the presentation of clinical symptoms that the user would have to identify and correlate with the imaging. This may provide a different view of correlating the physical examination and medical imaging in the context of a clinical setting. This investigation would explore if this specific simulation tool can be effectively adapted for use in teaching clinical cases. On a small scale this could be easily adapted into an advanced anatomy and imaging course that I teach for fourth year medical students to demonstrate a proof in concept prior to larger dissemination.

The ability to investigate the novice-expert paradigm within the context of simulation media has opened many additional avenues of research and collaboration in my future. With further investigation for larger scale use, this study demonstrates early implications for using simulation to teach the integration of medical imaging and the physical exam.

APPENDIX

Appendix 1: Clinical Interview Protocol

Equipment:

1. PAT simulation. Each user has a unique ID.
2. Digital audio recorder. Begins prior to start of interview and will stop when participant leaves the room.
3. Lamp. Room lights can't be dimmed.
4. Participant signs consent (sent to study coordinator prior to interview) and disclosure prior to beginning.

Introduction:

Welcome, and thank you for agreeing to participate in this activity. I am conducting a study about how medical students learn as part of my graduate research program. The purpose of this study is to find out how you think about imaging. I am going to ask you to think about some questions and tasks that you may never have even thought about before. That's completely fine. I am not at all interested in whether you get the right or wrong answers; the answer is basically irrelevant to my work. What I really need from you is for you to verbalize your thoughts as we go through the various tasks. The more you can talk out loud about your thinking the better it is for me. It will help me understand how you think about these things and how we can use it in education and patient care. There is no pressure and no grading here, I just want to hear your thoughts as you explore each task.

I am going to ask you questions, and will remind you to think aloud as you are going through each task.

I may make notes for myself as you talk about what you are thinking. But those notes are really just for me to flag interesting things that the audio might miss. I am not taking notes to evaluate you at all. It's just to remind me of interesting things that happen so I can find them when I go back later and look at the data.

Do you have any questions before we begin?

Recall – 5 minutes

You have had many opportunities to experience clinical care, and have had instances were patients had a CT scan as part of this clinical care.

In your clinical experiences, can you tell me about a case where you had a patient with imaging studies, specifically a CT scan. How did you use that information in patient care?

- Did you review the images? By yourself? With attending? With radiologist?
- Did you examine the patient? By yourself? With attending?
- Did you examine the patient, before or after the imaging?
- Did you correlate the imaging with the exam? How?

Knowledge construction – 20 minutes

Demonstrate how to use PAT.

A CT is projected within this mannequin and the use of this probe allows you to scroll through the image set as you scroll over the body in axial, sagittal and coronal planes. To scroll in the axial plane, you need to press the axial button on the iPad and then you are able to use the probe in axial plane by moving it in a superior and inferior direction, similarly for sagittal and coronal planes. We are only using axial and sagittal imaging planes for the cases I want you to look at. We will go through a couple of cases and if there is time at the end of our interview we can do additional cases. For now I would like you to get familiar with how the mechanics of the equipment work. Let me know when you are ready.

Cases:

For each case, you will be given an image on the iPad screen. I want you to think about it and please verbalize as you do so. Each image is located someplace in the body. I would like you to place the probe on the body where you think that slice is located in the body. When you are satisfied with your probe placement, push the 'ready' button on the iPad. At that time, the iPad image will become active and the case image will appear as a static image in the lower right corner of the screen. If they match to your satisfaction, push the 'complete' button on the iPad. If they don't match to your satisfaction, feel free to move the probe in any direction until the images match. When they match; then push the 'complete' button on the iPad.

[This task will be repeated for 6 cases (system now has unlimited presets). The user will get a test case to practice. The ideal would be to do 4 axial and 2 sagittal cases as a minimum.]

As you are going through the cases, please tell me what you are seeing and how you are making your decisions of where you are putting the probe.

[At the end of each case, the device provides a % success for that case on the screen. Will reassure user that if they do not see 100%, that it is okay. I will try to record these values during interview. These are not directly saved on backside, but calculated as a range.]

[The presets will be set to be nearly identical for each user.]

[Have put presets into simulator under heading of KC1,….KC15. Can choose from these images to test subjects]

[The device will record a time from when the participant sees the image (called Preset Selected) to the placement of the probe at the desired location (called Game Started). It will record the final placement of the probe (called Game Ended). Time will be collected from Preset selected to Game Start when user is first able to see the image, and time is continued to Game End. From Game Start to Game End, the movements of the probe and incremental time is recorded.]

Knowledge evaluation – 20 minutes

#1 When I worked with another student/faculty (novice/expert), they put the probe here (placing it on the body – will make a note if I placed it above or below the real location), to represent where this slice was located.
[3 presets …..KE1, KE2, and KE3]

- Do you agree with their probe placement? Why or why not?
- Why do you think they might have placed the probe there?
- If no, how would you help them understand why it is not there?
- Can you show me where you would put the probe and when you are ready, push 'ready' and then 'complete' like the previous cases?
- There will be 2 cases for this section –
    - 1st – wrong probe placement by another user – they will have to assess if that probe placement is correct or incorrect and then, show 'correct' location.
    - 2nd – probe placement too high or too low for a structure. They will have to assess if it is too high or too low to assess location of organ. For example: Interviewer places probe 2 cm below belly button. Then asks if this is too high or too low to assess the kidneys.

#2 The second preset case would be to a structure (stomach, duodenum, pancreas, kidneys, or something – will pick out a couple of organs and randomly use with group)
[6 presets …..KE2.1 to KE2.6]

- A student last semester said that the 'organ' was abnormal, why do you think they might have thought that?
- Do you agree?
    - If they agree, why do you think it is abnormal?
    - If they don't agree, why do you disagree? How could you explain to them that it is not abnormal?

- Would like to do 2 cases – one asking if the structure is normal and other asking it is abnormal. But using the framework that their peer called it normal or abnormal.

Scenario prediction/explanation – 15 minutes

#1 You are asked to consider a biopsy of this *'structure/thing'* and the patient has to remain supine. Because you are right handed, you choose to optimize your strengths and approach the patient from the left. *[Preset SP1 – sagittal image – so can predict structure out of image plane]*

- As you think about your approach, what structures or factors will you need to consider as you gain access to biopsy this *'thing'*?
- What would you expect to see as you 'scan'? (No movement of the probe)
    - Be specific! How would the image change as you scanned?

Now that you have had time to think about this, shall we see if that approach would encounter those items? (Allow them to move probe)

- As you do that, can you tell me what you are seeing?
- How successful would that have been?

#2 A situation will be presented where I was going to attempt to biopsy a lesion (in the liver), and present a very dangerous approach. Will ask participant if this is a reasonable and safe approach? If not, why? If yes, why? If they say no, then have them explain and show a better approach. *[These are axial images. They will see structures that they would encounter. Need to explain why it is safe or not.]*

[*Presets for a liver biopsy and renal biopsy, SP2, SP3, SP4. As I have it set up, I have small dots on the simulator that I will use as my entrance site for a biopsy needle. These access points would cross pleura or bowel.]*

Thank you participating in this study. Do you have any questions for me?
I would appreciate not sharing the activities of today with other classmates, as I want to get their unbiased impressions similar to your experience.
Give them their compensation gift: UW logo travel mug.

Total Interview Time: 60 minutes

Appendix 2: Data Collection Worksheet

Study ID			
Age		Gender	
Specialty choice		Yrs as Radiologist (faculty)	
Date		Audio Tape ID	

Recall	

Knowledge construction	Test case	axial				
				Preset ID	Final %	Content of slice
	case #1	axial	sagittal			
	case #2	axial	sagittal			
	case #3	axial	sagittal			
	case #4	axial	sagittal			
	case #5	axial	sagittal			
	case #6	axial	sagittal			

Knowledge evaluation	Wrong probe placement by other peer	axial	sagittal	Preset ID:	
	User assessment:				
	Probe placement by interviewer with respect to correct location	Probe placement with respect to preset:		Higher	Lower
		Preset ID:			
	User assessment:				
	Normal structure called 'Abnormal'	Axial	Organ:		
		Preset ID:			
	User Assessment:				
	Normal structure called 'Normal'	Axial	Organ:		
		Preset ID:			
	User Assessment:				

Scenario prediction	#1 Biopsy access/ expect to see (predict what they should see before viewing)				
		Location:	SP1		
		Orientation of image:	sagittal		
		Orientation of 'biopsy access':		axial	
	User Assessment:				
	#2 Show a potential access to biopsy site:			SP2 SP3 SP4	
		What organ being biopsied?			
		Is this is a reasonable approach:			
		Why or why not:			
		Can you plan a better approach?			
			Explain		
			Show		

References

Abstract Reference

Crowley, R. S., Naus, G. J., Stewart, J., & Friedman, C. P. (2003). Development of Visual Diagnostic Expertise in Pathology: An Information-processing Study. *Journal of the American Medical Informatics Association*, *10*(1), 39–51.

Chapter 1 References

Albanese, M. (2010). The gross anatomy laboratory: a prototype for simulation-based medical education. *Medical Education*, *44*(1), 7–9.

Albanese, M. A., Mejicano, G., Mullan, P., Kokotailo, P., & Gruppen, L. (2008). Defining characteristics of educational competencies. *Medical Education*, *42*(3), 248–255.

Asad, M. R., & Nasir, N. (2015). Role of Living and Surface Anatomy in Current Trends of Medical Education. *International Journal of Advance Research and Innovative Ideas in Education*, *1*(2), 203–210.

Azer, S. A. (2013). The place of surface anatomy in the medical literature and undergraduate anatomy textbooks. *Anatomical Sciences Education*, *6*(6), 415–432.

Bergman, E.M., Sieben, J.M., Smailbegovic, I., de Bruin, A.B.H., Scherpbier, A.J.J.A., & Van Der Vleuten, C.P.M. (2012). Constructive, collaborative, contextual, and self-directed learning in surface anatomy education. *Anatomical Sciences Education*, *6*(2), 114–124.

Bowen, J. L. (2006). Educational strategies to promote clinical diagnostic reasoning. *The New England Journal of Medicine*, *355*(21), 2217–2225.

Chen, F. M., Bauchner, H., & Burstin, H. (2004). A call for outcomes research in medical education. *Academic Medicine*, *79*(10), 955–960.

Collins, J. (2006). Medical Education Research: Challenges and Opportunities 1. *Radiology*, *240*(3), 639-647.

Cutler, D. M., & McClellan, M. (2001). Is Technological Change In Medicine Worth It? *Health Affairs*, *20*(5), 11–29.

Dauphinee, W. D., & Wood-Dauphinee, S. (2004). The need for evidence in medical education: the development of best evidence medical education as an opportunity to inform, guide, and sustain medical education research. *Academic Medicine*, *79*(10), 925–930.

Davis, J., Crabb, S., Rogers, E., Zamora, J., & Khan, K. (2008). Computer-based teaching is as good as face to face lecture-based teaching of evidence based medicine: a randomized controlled trial. *Medical Teacher*, *30*(3), 302–307.

Drake, R. L. (2007). A unique, innovative, and clinically oriented approach to anatomy education. *Academic Medicine*, *82*(5), 475–478.

Dusseau, J., Knutson, D., & Way, D. (2008). Anatomy correlations: introducing clinical skills to improve performance in anatomy. *Family Medicine*, *40*(9), 633–637.

Duszak, R., Jr. (2012). Medical imaging: is the growth boom over? [White paper]. Neiman Report, No.1, October 2012.

Ginsburg, H. (1981). The clinical interview in psychological research on mathematical thinking: Aims, rationales, techniques. *For the learning of mathematics*, *1*(3), 4-11.

Ginsburg, H. (1997). *Entering the child's mind: The clinical interview in psychological research and practice*. Cambridge: Cambridge University Press.

Granger, N. (2004). Dissection laboratory is vital to medical gross anatomy education. *The Anatomical Record*, 281, 6-8.

Grosvenor, L. J., Verma, R., O'Brien, R., Entwisle, J. J., & Finlay, D. (2003). Does Reporting of Plain Chest Radiographs Affect the Immediate Management of Patients Admitted to a Medical Assessment Unit? *Clinical Radiology*, *58*(9), 719–722.

Gunderman, R. B. R., & Wilson, P. K. P. (2005). Viewpoint: exploring the human interior: the roles of cadaver dissection and radiologic imaging in teaching anatomy. *Academic Medicine*, *80*(8), 745–749.

Guttmann, G., Drake, R., & Trelease, R. (2004). To what extent is cadaver dissection necessary to learn medical gross anatomy? A debate forum. The Anatomical Record, 281, 2-3.

Hardy, M., Snaith, B., & Scally, A. (2013). The impact of immediate reporting on interpretive discrepancies and patient referral pathways within the emergency department: a randomised controlled trial. *The British Journal of Radiology*, *86*, 20120112.

Hyde, L., Erolin, C., & Ker, J. (2012). Creation of abdominal palpation model prototype for training of medical students in detection and diagnosis of liver disease. *Journal of Visual Communication in Medicine*, *35*(3), 104–114.

Issenberg, B.S., McGaghie, W. C., Petrusa, E. R., Lee Gordon, D., & Scalese, R. J. (2005). Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Medical Teacher*, *27*(1), 10–28.

Jenkinson, J. (2009). Measuring the Effectiveness of Educational Technology: what are we attempting to Measure. *Electronic Journal of E-Learning*, *7*(3), 273–280.

Jones, N. A., Olafson, R. P., & Sutin, J. (1978). Evaluation of a gross anatomy program without dissection. *Journal of Medical Education*, *53*(3), 198–205.

Kachlik, D., Baca, V., Bozdechova, I., Cech, P., & Musil, V. (2008). Anatomical terminology and nomenclature: past, present and highlights. *Surgical and Radiologic Anatomy*, *30*(6), 459–466.

Kiu, A., Bano, F., Barnes, R., & Khan, S. H. M. (2010). IRMER regulations: compliance rate of radiograph reporting by non-radiology clinicians. *Clinical Radiology*, *65*(12), 984–988.

Kotzé, S. H., Driescher, N. D., & Mole, C. G. (2013). The translucent cadaver: A follow-up study to gauge the efficacy of implementing changes suggested by students. *Anatomical Sciences Education*, *6*(6), 433–439.

Kotzé, S. H., Mole, C. G., & Greyling, L. M. (2012). The translucent cadaver: An evaluation of the use of full body digital x-ray images and drawings in surface anatomy education. *Anatomical Sciences Education*, *5*(5), 287–294.

Lee, K., Whelan, J. S., Tannery, N. H., Kanter, S. L., & Peters, A. S. (2013). 50 years of publication in the field of medical education. *Medical teacher*, 35(7), 591-598.

Lichtenberg, F. R. (2011). The quality of medical care, behavioral risk factors, and longevity growth. *International Journal of Health Care Finance and Economics*, *11*(1), 1–34.

Mahant, S., Jovcevska, V., & Wadhwa, A. (2012). The Nature of Excellent Clinicians at an Academic Health Science Center. *Academic Medicine*, *87*(12), 1715–1721.

Maudsley, G. (2011). Mixing it but not mixed-up: mixed methods research in medical education (a critical narrative review). *Medical teacher*, *33*(2), e92-e104.

McLachlan, J. C., & De Bere, S. R. (2004). How we teach anatomy without cadavers. *The Clinical Teacher*, 1(2), 49-52.

Miller, G. E. (1990). The assessment of clinical skills/competence/performance. *Academic Medicine*, *65*(9 Suppl), S63–7.

O'Brien, K. E., Cannarozzi, M. L., Torre, D. M., Mechaber, A. J., & Durning, S. J. (2008). Training and Assessment of CXR/Basic Radiology Interpretation Skills: Results From the 2005 CDIM Survey. *Teaching and Learning in Medicine*, *20*(2), 157–162.

Pasquale, S. J. (2013). Education and Learning Theory. In Levine, A.I, DeMaria, S., Schwartz, A.D. & Sim, A.J. (Eds.), *The Comprehensive Textbook of Healthcare Simulation* (pp. 51–55). New York: Springer Science & Business Media.

Pathiraja, F., Little, D., & Denison, A. R. (2014). Are radiologists the contemporary anatomists? *Clinical Radiology*, *69*(5), 458–461.

Patten, D. (2007). What lies beneath: the in living anatomy teaching. *The Clinical Teacher*, *4*(1), 10–14.

Posner, G. J., & Gertzog, W. A. (1982). The clinical interview and the measurement of conceptual change. *Science Education*, *66*(2), 195-209.

Rector, A. L. (1999). Clinical terminology: why is it so hard? *Methods of Information in Medicine*, *38*(4-5), 239–252.

Regehr, G. (2004). Trends in medical education research. *Academic Medicine*, *79*(10), 939–947.

Rizzolo, L. J., Rando, W. C., O'Brien, M. K., Haims, A. H., Abrahams, J. J., & Stewart, W. B. (2010). Design, implementation, and evaluation of an innovative anatomy course. *Anatomical Sciences Education*, *3*(3), 109–120.

Rubin, A. (1989). Contributions of Cognitive Science and Educational Technology to Training in Radiology. *Investigative Radiology*, *24*, 729–732.

Schaefer, J. J., III, Vanderbilt, A. A., Cason, C. L., Bauman, E. B., Glavin, R. J., Lee, F. W., & Navedo, D. D. (2011). Literature Review. *Simulation in Healthcare: the Journal of the Society for Simulation in Healthcare*, *6*, S30–S41.

Scheiner, J. D., Noto, R. B., & McCarten, K. M. (2002). Importance of radiology clerkships in teaching medical students life-threatening abnormalities on conventional chest radiographs. *Academic Radiology*, *9*(2), 217–220.

Seagull, F. J., Bailey, J. E., Trout, A., Cohan, R. H., & Lypson, M. L. (2014). Residents' Ability to Interpret Radiology Images: Development and improvement of an assessment tool. *Academic Radiology*, *21*(7), 909–915.

Sugand, K., Abrahams, P., & Khurana, A. (2010). The anatomy of anatomy: A review for its modernization. *Anatomical Sciences Education*, 3(2), 83-93.

Teddlie, C. & Tashakkori, A. (2011). Mixed methods research: Contemporary issues in an emerging field. In Denzin, N. K., & Lincoln, Y. S. (Eds.), *The Sage handbook of qualitative research* (pp. 285-299). Los Angeles: Sage Publications.

Tyler, R. W. (1949). *Basic principles of curriculum and instruction*. University of Chicago press.

Verghese, A., Brady, E., Kapur, C. C., & Horwitz, R. I. (2011). The bedside evaluation: ritual and reason. *Annals of Internal Medicine*, *155*(8), 550–553.

Wenrich, M. D., Jackson, M. B., Ajam, K. S., Wolfhagen, I. H., Ramsey, P. G., & Scherpbier, A. J. (2011). Teachers as Learners: The Effect of Bedside Teaching on the Clinical Skills of Clinician–Teachers. *Academic Medicine*, *86*(7), 846–852.

Wilson, D. R., & Nava, P. B. (2010). Medical student responses to clinical procedure teaching in the anatomy lab. *The Clinical Teacher*, 7(1), 14–18.

Chapter 2 References

Afaq, A., & McCall, J. (2002). Improving undergraduate education in radiology. *Academic Radiology*, *9*(2), 221–223.

Albanese M. (2010). The gross anatomy laboratory: a prototype for simulation-based medical education. *Medical Education*, *44*(1), 7–9.

Allen, I.E. & Seaman, C.A. (2007). Likert scales and data analyses. *Quality Progress, 40* (7), 64-65.

Asad, M.R. & Nasir, N. (2015). Role of Living and Surface Anatomy in Current Trends of Medical Education. *International Journal of Advance Research and Innovative Ideas in Education*, *1*, (2), 203–210.

Azer, S.A. (2013). The place of surface anatomy in the medical literature and undergraduate anatomy textbooks. *Anatomical Sciences Education*, *6*(6), 415–432.

Bergman, E.M., Sieben, J.M., Smailbegovic, I., de Bruin, A.B.H., Scherpbier, A.J.J.A., & Van Der Vleuten, C.P.M. (2012). Constructive, collaborative, contextual, and self-directed learning in surface anatomy education. *Anatomical Sciences Education*, *6*(2), 114–124.

Bradley, P. (2003). Simulation in clinical learning. *Medical Education*, *37*(suppl 1), 1–5.

Bradley, P. (2006). The history of simulation in medical education and possible future directions. *Medical Education*, *40*(3), 254–262.

Branstetter, B. F., IV, Faix, L. E., Humphrey, A. L., & Schumann, J. B. (2007). Preclinical Medical Student Training in Radiology: The Effect of Early Exposure. *American Journal of Roentgenology*, *188*(1), W9–W14.

Chauvin, S. W. (2012). Educational Principles in Simulation. *Simulation in Radiology,* pp. 47-65.

Collins, J., Dotti, S. L., & Albanese, M. A. (2002). Teaching radiology to medical students: an integrated approach. *Academic Radiology*, *9*(9), 1046–1053.

Cooke, M., Irby, D.M., & O'Brien, B.C. (2010). *Educating physicians: a call for reform of medical school and residency*. San Francisco: Jossey-Bass.

Dillon, J. E., & Slanetz, P. J. (2010). Teaching Evidence-Based Imaging in the Radiology Clerkship Using the ACR Appropriateness Criteria. *Academic Radiology*, *17*(7), 912–916.

Dmytriw, A. A., Mok, P. S., Gorelik, N., Kavanaugh, J., & Brown, P. (2015). Radiology in the Undergraduate Medical Curriculum: Too Little, Too Late? *Medical Science Educator*, *25*(3), 223–227.

Drake, R.L. (2007). A unique, innovative, and clinically oriented approach to anatomy education. *Academic Medicine*, *82*(5), 475–478.

Dusseau, J., Knutson, D., & Way, D. (2008). Anatomy correlations: introducing clinical skills to improve performance in anatomy. *Family Medicine*, *40*(9), 633–637.

Duvivier, R. J., van Dalen, J., Muijtjens, A. M., Moulaert, V. R., van der Vleuten, C. P., & Scherpbier, A. J. (2011). The role of deliberate practice in the acquisition of clinical skills. *BMC Medical Education*, *11*(1), 101.

Eisen, L. A., Berger, J. S., Hegde, A., & Schneider, R. F. (2006). Competency in chest radiography. *Journal of General Internal Medicine*, *21*(5), 460–465.

Ericsson, K. A. (2004). Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Academic Medicine*, *79*(10 Suppl), S70–81.

Ericsson, K. A. (2006). Protocol analysis and expert thought: Concurrent verbalizations of thinking during experts' performance on representative tasks. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge Handbook of Expertise and Expert Performance* (pp. 223–241). New York: Cambridge University Press.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review, 87*(3), 215–51.

Feigin, D. S., Magid, D., Smirniotopoulos, J. G., & Carbognin, S. J. (2007). Learning and Retaining Normal Radiographic Chest Anatomy. *Academic Radiology*, *14*(9), 1137–1142.

Fonteyn, M. E., Kuipers, B., & Grobe, S. J. (1993). A description of think aloud method and protocol analysis. *Qualitative Health Research*, *3*(4), 430–441.

Friedman, C.P. (1995). Anatomy of the clinical simulation. *Academic Medicine*, *70*(3), 205–209.

Games Learning Society (2012). ADAGE [Computer software]. Madison, WI: University of Wisconsin–Madison. Retrieved from: http://www.gameslearningsociety.org/

Ganske, I., Su, T., Loukas, M., & Shaffer, K. (2006). Teaching Methods in Anatomy Courses in North American Medical Schools. *Academic Radiology*, *13*(8), 1038–1046.

Garland, R. (1991). The mid-point on a rating scale: Is it desirable? *Marketing Bulletin, 2*(3), 66-70.

Granger, N.A. (2004). Dissection laboratory is vital to medical gross anatomy education. *The Anatomical Record Part B: The New Anatomist*, *281*(1), 6-8.

Gunderman, R.B.R., & Wilson, P.K.P. (2005). Viewpoint: exploring the human interior: the roles of cadaver dissection and radiologic imaging in teaching anatomy. *Academic Medicine*, *80*(8), 745–749.

Guttmann, G., Drake, R., & Trelease, R. (2004). To what extent is cadaver dissection necessary to learn medical gross anatomy? A debate forum. *The Anatomical Record*, *281*, 2-3.

Hall, S. How to use the Likert scale in statistical analysis. http://www.ehow.com/how_4855078_use-likert-scale-statistical-analysis.html. Accessed 12/13/2014. September 13, 2014.

Hyde, L., Erolin, C., & Ker, J. (2012). Creation of abdominal palpation model prototype for training of medical students in detection and diagnosis of liver disease. *Journal of Visual Communication in Medicine*, *35*(3), 104–114.

Issenberg, B.S., McGaghie, W.C., Petrusa, E.R., Lee Gordon, D., Scalese, R.J. (2005). Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Medical Teacher*, *27*(1), 10–28.

Jeffrey, D. R., Goddard, P. R., Callaway, M. P., & Greenwood, R. (2003). Chest Radiograph Interpretation by Medical Students. *Clinical Radiology*, *58*(6), 478–481.

Jenkinson, J. (2009). Measuring the Effectiveness of Educational Technology: what are we attempting to Measure. *Electronic Journal of E-Learning*, *7*(3), 273–280.

Johnson, E.O., Charchanti, A.V., & Troupis, T.G. (2012). Modernization of an anatomy class: From conceptualization to implementation. A case for integrated multimodal-multidisciplinary teaching. *Anatomical Sciences Education, 5*(6), 354–366.

Jones, N.A., Olafson, R.P., & Sutin, J. (1978). Evaluation of a gross anatomy program without dissection. *Journal of Medical Education*, *53*(3), 198–205.

Kalet, A. L., Song, H. S., Sarpel, U., Schwartz, R., Brenner, J., Ark, T. K., & Plass, J. (2012). Just enough, but not too much interactivity leads to better clinical skills performance after a computer assisted learning module. *Medical Teacher*, *34*(10), 833–839.

Lane, J. L., Slavin, S., & Ziv, A. (2001). Simulation in medical education: A review. *Simulation & Gaming*, *32* (3), 297-314.

Leschied, J. R., Knoepp, U. S., Hoff, C. N., Mazza, M. B., Klein, K. A., Mullan, P. B., & Kelly, A. M. (2013). Emergency Radiology Elective Improves Second-Year Medical Students' Perceived Confidence and Knowledge of Appropriate Imaging Utilization. *Academic Radiology*, *20*(9), 1168–1176.

Lewis, P., & Shaffer, K. (2005). Developing a national medical student curriculum in radiology. *Journal of the American College of Radiology, 2*(1), 8-11.

Lewis, P., & Shaffer, K. (2010). AMSER national medical student curriculum in radiology (out of print). *MedEdPORTAL Publications*, *6*, 7897.

Manning, D. (2010). Cognitive factors in reading medical images: A survey of cognitive factors and models of medical image interpretation. In Samei, E. & Krupinski, E. (Eds.) *The Handbook of Medical Image and Perception Techniques*, (pp.91-106), New York: Cambridge University Press.

Maxwell, J. A. (2004). Causal explanation, qualitative research, and scientific inquiry in education. *Educational Researcher*, *33*(2), 3-11.

McGaghie, W. C., Issenberg, S. B., Petrusa, E. R., & Scalese, R. J. (2006). Effect of practice on standardized learning outcomes in simulation-based medical education. *Medical Education*, *40*(8), 792–797.

Merriam, S.B., Caffarella, R.S., & Baumgartner, L.M. (2007). Traditional Learning Theories. In *Learning in Adulthood: A comprehensive guide*, (3 ed.). (pp.275-297). San Francisco: John Wiley & Sons.

Mirsadraee, S., Mankad, K., McCoubrie, P., Roberts, T., & Kessel, D. (2012). Radiology curriculum for undergraduate medical studies. *Clinical Radiology*, *67*(12), 1155–1161.

Murray, D., Boulet, J., Ziv, A., Woodhouse, J., Kras, J., & McAllister, J. (2002). An acute care skills evaluation for graduating medical students: a pilot study using clinical simulation. *Medical Education*, *36*(9), 833–841.

Owen, V.E., Ramirez, D., Salmon, A. & Halverson, R. (2014, April). *Capturing Learner Trajectories in Educational Games through ADAGE (Assessment Data Aggregator for Game Environments): A Click-Stream Data Framework for Assessment of Learning in Play.* Presented at the 2014 American Educational Research Association Annual Meeting, Philadelphia, PA.

Pasquale, S. J. (2013). Education and Learning Theory. In *The Comprehensive Textbook of Healthcare Simulation* (pp. 51–55). New York: Springer Science & Business Media.

Pathiraja, F., Little, D., & Denison, A.R. (2014). Are radiologists the contemporary anatomists? *Clinical Radiology*, *69*(5), 458–461.

Phillips, A. W., Smith, S. G., & Straus, C. M. (2013). The Role of Radiology in Preclinical Anatomy:. *Academic Radiology*, *20*(3), 297–304.e1.

Polhemus: Innovation in Motion. (2014). Selecting the Optimal Motion Tracker for Medical Training Simulators. Colchester, VT.

Reddy, S., Straus, C. M., McNulty, N. J., Tigges, S., Ayoob, A., Randazzo, W., Neutze, J., & Lewis, P. (2015). Development of the AMSER Standardized Examinations in Radiology for Medical Students. *Academic radiology*, *22*(1), 130-134.

Rizzolo, L.J., Rando, W.C., O'Brien, M.K., Haims, A.H., Abrahams, J.J., & Stewart, W.B. (2010). Design, implementation, and evaluation of an innovative anatomy course. *Anatomical Sciences Education*, *3*(3), 109–120.

Savran MM, Tranum-Jensen J, Clementsen PF, Svendsen JH, Pedersen JH, Poulsen SS, Arendrup, H., & Konge, L. (2015). Are medical students being taught anatomy in a way that best prepares them to be a physician? *Clinical Anatomy*, *28*(5), 568–575.

Scheiner, J. D., Noto, R. B., & McCarten, K. M. (2002). Importance of radiology clerkships in teaching medical students life-threatening abnormalities on conventional chest radiographs. *Academic Radiology*, *9*(2), 217–220.

Schumacher, D.J., Englander, R., & Carraccio, C. (2013). Developing the Master Learner. *Academic Medicine*, *88*(11), 1635–1645.

Straus, C. M., Webb, E. M., Kondo, K. L., Phillips, A. W., Naeger, D. M., Carrico, C. W., et al. (2014). Medical student radiology education: summary and recommendations from a national survey of medical school and radiology department leadership. *Journal of the American College of Radiology : JACR*, *11*(6), 606–610.

Sugand, K., Abrahams, P., & Khurana, A. (2010). The anatomy of anatomy: A review for its modernization. *Anatomical Sciences Education*, *3*(2), 83-93.

Verghese, A., Brady, E., Kapur, C.C., & Horwitz, R.I. (2011). The bedside evaluation: ritual and reason. *Annals of Internal Medicine*, *155*(8), 550–553.

Walker, D. & Myrick, F. (2006). Grounded Theory: An Exploration of Process and Procedure. *Qualitative Health Research*, *16*(4), 547–559.

Wenrich, M.D., Jackson, M.B., Ajam, K.S., Wolfhagen, I.H., Ramsey, P.G., & Scherpbier, A.J. (2011). Teachers as Learners: The Effect of Bedside Teaching on the Clinical Skills of Clinician–Teachers. *Academic Medicine*, *86*(7), 846–852.

Wilson, D.R., & Nava, P.B. (2010). Medical student responses to clinical procedure teaching in the anatomy lab. *The Clinical Teacher*, *7*(1), 14–18.

Chapter 3 References

Applefield, J. M., Huber, R., & Moallem, M. (2000). Constructivism in theory and practice: Toward a better understanding. *The High School Journal*, *84*(2), 35–53.

Atkinson, R., & Flint, J. (2001). Accessing hidden and hard-to-reach populations: Snowball research strategies. *Social research update*, *33*(1), 1-4.

Azevedo, R. (1999). Expertise in Radiology: Accounting for the Evidence and Implications for Instruction. Paper presented at the Annual Meeting of the American Educational Research Association (AERA). Montreal, Ontario, Canada.

Berbaum, K. S., Franken, E. A., Dorfman, D. D., Rooholamini, S. A., Kathol, M. H., Barloon, T. J., … Montgomery, W.J. (1990). Satisfaction of search in diagnostic radiology. *Investigative Radiology*, *25*(2), 133–140.

Biernacki, P. & Waldorf, D. (1981). Snowball Sampling: Problems and techniques of chain referral sampling. *Sociological Methods & Research, 10*(2), 141-163.

Bodner, G. M. (1986). Constructivism: A theory of knowledge. *Journal of Chemical Education, 63* (10), 873-878.

Bordage, G. (2007). Prototypes and semantic qualifiers: from past to present. Medical Education, 41(12), 1117–1121.

Bowen, J. L. (2006). Educational strategies to promote clinical diagnostic reasoning. *The New England Journal of Medicine*, *355*(21), 2217–2225.

Bransford, J.D., Brown, A.L., & Cocking, R.R. (2000). *How People Learn: Brain, Mind, Experience and School, Expanded Edition*. Washington, DC: National Academy Press.

Campbell, I. (2007). Chi-squared and Fisher-Irwin tests of two-by-two tables with small sample recommendations. *Statistics in Medicine, 26*, 3661-3675.

Chi, M. T. H. (2006a). Laboratory methods for assessing experts' and novices' knowledge. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge Handbook of Expertise and Expert Performance* (pp. 167-184). New York: Cambridge University Press.

Chi, M. T. H. (2006b). Two approaches to the study of experts' characteristics. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge Handbook of Expertise and Expert Performance* (pp. 21–30). New York: Cambridge University Press.

Chi, M. T. H. (2011). Theoretical Perspectives, Methodological Approaches, and Trends in the Study of Expertise. In Y. Li & G. Kaiser (Eds.), *Expertise in Mathematical Instruction* (pp. 17–39). Boston, MA: Springer US.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and Representation of Physics Problems by Experts and Novices. *Cognitive Science, 5*(2), 121–152.

Cohen, L., Manion, L., & Morrison, K. (2007). Sampling. In *Research Methods in Education, 6th Ed.* (pp. 100-118). New York, NY: Routledge.

Crowley, R. S., Naus, G. J., & Friedman, C. P. (2001). Development of visual diagnostic expertise in pathology. In *Proceedings of the AMIA Symposium* (p. 125). American Medical Informatics Association.

Crowley, R. S., Naus, G. J., Stewart, J., & Friedman, C. P. (2003). Development of Visual Diagnostic Expertise in Pathology: An Information-processing Study. *Journal of the American Medical Informatics Association*, *10*(1), 39–51.

Cutler, D. M., & McClellan, M. (2001). Is Technological Change In Medicine Worth It? *Health Affairs, 20*(5), 11–29.

Daley, B. J. (2000). Learning in professional practice. *New Directions for Adult and Continuing Education*, *2000*(86), 33-42.

de Bruin, A. B. H., Schmidt, H. G., & Rikers, R. M. J. P. (2005). The role of basic science knowledge and clinical knowledge in diagnostic reasoning: a structural equation modeling approach. *Academic Medicine, 80*(8), 765–773.

Dreyfus, S. E. (2004). The five-stage model of adult skill acquisition. *Bulletin of science, technology & society*, *24*(3), 177-181.

Dreyfus, S. E., & Dreyfus, H. L. (1980). A five-stage model of the mental activities involved in directed skill acquisition (No. ORC-80-2). California Univ Berkeley Operations Research Center.

Dunphy, B. C., & Williamson, S. L. (2004). In pursuit of expertise. Toward an educational model for expertise development. *Advances in Health Sciences Education : Theory and Practice*, *9*(2), 107–127.

Duszak, R., Jr. (2012). Medical imaging: is the growth boom over. Neiman Report. Brief.

Ericsson, K. A. (1996). The acquisition of expert performance: An introduction to some of the issues. In K. A. Ericsson (Ed.), *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games* (pp. 1-50). Mahwah, NJ: Lawrence Erlbaum.

Ericsson, K. A. (2004). Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Academic Medicine*, *79*(10 Suppl), S70–81.

Ericsson, K. A. (2006). Protocol analysis and expert thought: Concurrent verbalizations of thinking during experts' performance on representative tasks. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge Handbook of Expertise and Expert Performance* (pp. 223–241). New York: Cambridge University Press.

Ericsson, K. A. (2007). An expert-performance perspective of research on medical expertise: the study of clinical performance. *Medical Education*, *41*(12), 1124–1130.

Ericsson, K. A. (2008). Deliberate Practice and Acquisition of Expert Performance: A General Overview. *Academic Emergency Medicine*, *15*(11), 988–994.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*(3), 215–251.

Faruque, J., Rubin, D. L., Beaulieu, C. F., & Napel, S. (2012). Modeling Perceptual Similarity Measures in CT Images of Focal Liver Lesions. *Journal of Digital Imaging*, *26*(4), 714–720.

Ferguson, K.J. (2010). Facilitating student learning. In W.B. Jeffries & K.N. Huggett (Eds.), *An Introduction to Medical Teaching* (pp. 1- 10). New York: Springer.

Finnerty, E. P., Chauvin, S., Bonaminio, G., Andrews, M., Carroll, R. G., & Pangaro, L. N. (2010). Flexner revisited: the role and value of the basic sciences in medical education. *Academic Medicine*, *85*(2), 349-355.

Fonteyn, M. E., Kuipers, B., & Grobe, S. J. (1993). A description of think aloud method and protocol analysis. *Qualitative Health Research*, *3*(4), 430–441.

Getty, D. J., Pickett, R. M., D'Orsi, C. J., & Swets, J. A. (1988). Enhanced interpretation of diagnostic images. *Investigative Radiology, 23*(4), 240–252.

Glaser, R. (1984). Education and thinking: The role of knowledge. *American psychologist*, *39*(2), 93.

Glaser, R., & Chi, M.T.H. (1988) Overview. In Chi, M. T., Glaser, R., & Farr, M. J. (Eds.) *The nature of expertise* (pp. xv-xxviii). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gonzalez, C., & Thomas, R. P. (2008). Effects of Automatic Detection on Dynamic Decision Making. *Journal of Cognitive Engineering and Decision Making*, *2*(4), 328–348.

Gould, J. D. (1978). How experts dictate. *Journal of Experimental Psychology: Human Perception and Performance, 4*, 648–661.

Grill-Spector, K., & Kanwisher, N. (2005). Visual recognition as soon as you know it is there, you know what it is. *Psychological Science*, *16*(2), 152-160.

Grosswald, S. J. (1992). Problem-solving strategies of experienced and novice physicians. *Journal of Continuing Education in the Health Professions, 12,* 205-213.

Kalet, A. L., Song, H. S., Sarpel, U., Schwartz, R., Brenner, J., Ark, T. K., & Plass, J. (2012). Just enough, but not too much interactivity leads to better clinical skills performance after a computer assisted learning module. *Medical Teacher*, *34*(10), 833–839.

Kok, E. M., de Bruin, A. B. H., Robben, S. G. F., & van Merriënboer, J. J. G. (2012). Looking in the Same Manner but Seeing it Differently: Bottom-up and Expertise Effects in Radiology. *Applied Cognitive Psychology*, *26*(6), 854–862.

Krupinski, E. A. (1996). Visual scanning patterns of radiologists searching mammograms. *Academic Radiology*, *3*(2), 137–144.

Krupinski, E. A. (2000a). The importance of perception research in medical imaging. *Radiation Medicine*, *18*(6), 329–334.

Krupinski, E. A. (2000b). Medical image perception: evaluating the role of experience (Vol. 3959, pp. 281–289). Presented at the Proceedings of the SPIE.

Krupinski, E. A., Berger, W. G., Dallas, W. J., & Roehrig, H. (2003). Searching for nodules: What features attract attention and influence detection? *Academic Radiology*, *10*(8), 861–868.

Kumazawa, S., Muramatsu, C., Li, Q., Li, F., Shiraishi, J., Caligiuri, P., Schmidt, R.A., MacMahon, H., & Doi, K. (2008). An Investigation of Radiologists' Perception of Lesion Similarity. *Academic Radiology*, *15*(7), 887–894.

Kundel, H. L. (2006). History of Research in Medical Image Perception. *Journal of the American College of Radiology*, *3*(6), 402–408.

Kundel, H. L., & La Follette, P. S. (1972). Visual search patterns and experience with radiological images. *Radiology*, *103*(3), 523–528.

Kundel, H. L., Nodine, C. F., & Krupinski, E. A. (1989). Searching for Lung Nodules - Visual Dwell Indicates Locations of False-Positive and False-Negative Decisions. *Investigative Radiology*, *24*(6), 472–478.

Laerd Statistics (2015). Statistical tutorials and software guides. Retrieved from https://statistics.laerd.com/

Lesgold, A. M. (1983). Acquiring Expertise (pp. 1–51). Office of Naval Research.

Lesgold, A. M., Feltovich, P. J., Glaser, R., & Wang, Y. (1981). The Acquisition of Perceptual Diagnostic Skill in Radiology (pp. 1–155). Office of Naval Research.

Lesgold, A., Rubinson, H., Feltovich, P., Glaser, R., Klopfer, D., & Wang, Y. (1988) Expertise in a Complex Skill: Diagnosing X-Ray Pictures. In Chi, M. T., Glaser, R., & Farr, M. J. (Eds.) *The nature of expertise* (pp. 311–342). Hillsdale, NJ: Lawrence Erlbaum Associates.

Mahant, S., Jovcevska, V., & Wadhwa, A. (2012). The Nature of Excellent Clinicians at an Academic Health Science Center. *Academic Medicine*, *87*(12), 1715–1721.

Mann, K. V. (2010). Theoretical perspectives in medical education: past experience and future possibilities. *Medical Education*, *45*(1), 60–68.

Mehmood, I., Ejaz, N., Sajjad, M., & Baik, S. W. (2013). Computers in Biology and Medicine. *Computers in Biology and Medicine*, *43*(10), 1471–1483.

Mello-Thoms, C. (2003). Perception of breast cancer: Eye-position analysis of mammogram interpretation. *Academic Radiology*, *10*(1), 4–12.

Mello-Thoms, C. (2006). How Does the Perception of a Lesion Influence Visual Search Strategy in Mammogram Reading? *Academic Radiology*, *13*(3), 275–288.

Mello-Thoms, C., Dunn, S., Nodine, C. F., Kundel, H. L., & Weinstein, S. P. (2002). The perception of breast cancer: What differentiates missed from reported cancers in mammography? *Academic Radiology*, *9*(9), 1004–1012.

Mello-Thoms, C., & Hardesty, L. (2005). How does lesion conspicuity affect visual search strategy in mammogram reading? (Vol. 5749, pp. 86–93). Presented at the Proceedings of the SPIE.

Merriam, S.B., Caffarella, R.S., & Baumgartner, L.M. (2007). Traditional Learning Theories. In *Learning in Adulthood: A comprehensive guide*, (3 ed.). (pp.275-297). San Francisco: John Wiley & Sons.

Miller, G. A. (1956). The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychological Review, 63*(2), 81–97.

Nendaz, M. R., & Bordage, G. (2002). Promoting diagnostic problem representation. *Medical Education, 36*(8), 760–766.

Nodine, C. F., & Krupinski, E. A. (1998). Perceptual skill, radiology expertise, and visual test performance with NINA and WALDO. *Academic Radiology, 5*, 603-612.

Nodine, C. F., & Mello-Thoms, C. (2010). The role of expertise in radiologic image interpretation. In E. Samei & E. Krupinski (Eds.), *The Handbook of Medical Image Perception and Techniques* (pp. 139–156). Cambridge University Press New York, NY.

Norman, G. (2005). Research in clinical reasoning: past history and current trends. *Medical Education*, *39*(4), 418–427.

Norman, G., Eva, K., Brooks, L., & Hamstra, S. (2006). Expertise in Medicine and Surgery. In Ericsson, K.A., Charness, N., Feltovich, P.J., & Hoffman, R.R., *The Cambridge handbook of Expertise and Expert Performance.* (pp. 339-353). New York: Cambridge University Press.

Pathiraja, F., Little, D., & Denison, A. R. (2014). Are radiologists the contemporary anatomists? *Clinical Radiology, 69*(5), 458–461.

Piaget, J. (1929). The concept of "life." In *The child's conception of the world.* (pp. 194–206). London: Routledge and Kegan Paul, Ltd.

Piaget, J. (1964). Part I: Cognitive development in children: Piaget development and learning. *Journal of Research in Science Teaching, 2*(3), 176–186.

Richardson, J.T.E. (2011). The analysis of 2 x 2 contingency tables - Yet again. *Statistics in Medicine,* 30, 890.

Rikers, R.M.J.P., Loyens, S.M.M. & Schmidt, H.G. (2004). The role of encapsulated knowledge in clinical case representations of medical students and family doctors. *Medical Education, 38*, 1035-1043.

Rubin, A. (1989). Contributions of Cognitive Science and Educational Technology to Training in Radiology. Investigative *Radiology, 24*, 729–732.

Schmidt, H. G., & Boshuizen, H. P. A. (1993). On acquiring expertise in medicine. *Educational Psychology Review*, *5*, 205–221.

Schmidt, H. G., & Rikers, R. M. J. P. (2007). How expertise develops in medicine: knowledge encapsulation and illness script formation. *Medical Education, 41*, 1133-1139.

Schunk, D. H. (2012). Constructivism. In *Learning Theories* (6 ed.). (pp. 228-277). Boston: Pearson.

Stevens, C., Winskel, H., Howell, C., Vidal, L.-M., Latimer, C., & Milne-Home, J. (2010). Perceiving dance: schematic expectations guide experts' scanning of a contemporary dance film. *Journal of Dance Medicine & Science: Official Publication of the International Association for Dance Medicine & Science*, *14*(1), 19–25.

van de Wiel, M. W., Schmidt, H. G., & Boshuizen, H. P. (1998). A failure to reproduce the intermediate effect in clinical case recall. *Academic Medicine*, *73*(8), 894–900.

Vanderbilt, A. A., Feldman, M., & Wood, I. K. (2013). Assessment in undergraduate medical education: a review of course exams. *Medical Education Online*, *18*(0), 368. http://doi.org/10.1002/chp.20001

VanSickle, R. L., & Hoge, J. D. (1991). Higher cognitive thinking skills in social studies: Concepts and critiques. *Theory & Research in Social Education, 19*(2), 152–172.

Voisin, S., Yoon, H.J., Tourassi, G., Morin-Ducote, G., & Hudson, K. (2013). Personalized modeling of human gaze: Exploratory investigation on mammogram readings (pp. 1–4). Presented at the Biomedical Sciences and Engineering Conference (BSEC), 2013, IEEE.

Chapter 4 References

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, *89*(4), 369-406.

Anderson, J. R., & Bower, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review, 79*(2), 97-123.

Arnold, L., Willoughby, L., Calkins, V., & Gammon, L. (1981). Use of peer evaluation in the assessment of medical students. *Journal of Medical Education, 56*, 35–42.

Arzi, H. J., Ben-Zvi, R., & Ganiel, U. (1986). Forgetting versus savings: The many facets of long-term retention. *Science Education, 70*(2), 171–188.

Azevedo, R., Faremo, S., & Lajoie, S. P. (2007, January). Expert-novice differences in mammogram interpretation. In *Proceedings of the Cognitive Science Society* (Vol. 29, No. 29), 65-70.

Barz, D. L., & Achimaş-Cadariu, A. (2016). The Development of Scientific Reasoning in Medical Education: A Psychological Perspective. *Clujul Medical*, *89*(1), 32-37.

Bowen, J. L. (2006). Educational strategies to promote clinical diagnostic reasoning. *New England Journal of Medicine*, *355*(21), 2217–2225.

Bransford, J.D., Brown, A.L., & Cocking, R.R. (2000). *How People Learn: Brain, Mind, Experience and School, Expanded Edition*. Washington, DC: National Academy Press.

Brandt, K. R., Cooper, L. M., & Dewhurst, S. A. (2005). Expertise and recollective experience: recognition memory for familiar and unfamiliar academic subjects. *Applied Cognitive Psychology*, *19*(9), 1113–1125.

Chauvin, S.W. (1988) Educational Principles in Simulation. In Chi, M. T., Glaser, R., & Farr, M. J. (Eds.) *The nature of expertise* (pp. 47-65). Hillsdale, NJ: Lawrence Erlbaum Associates.

Chi, M. T. H. (2006a). Laboratory methods for assessing experts' and novices' knowledge. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge Handbook of Expertise and Expert Performance* (pp. 167-184). New York: Cambridge University Press.

Chi, M. T. H. (2006b). Two approaches to the study of experts' characteristics. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge Handbook of Expertise and Expert Performance* (pp. 21–30). New York: Cambridge University Press.

Chi, M. T. H. (2011). Theoretical Perspectives, Methodological Approaches, and Trends in the Study of Expertise. In Y. Li & G. Kaiser (Eds.), *Expertise in Mathematical Instruction* (pp. 17–39). Boston, MA: Springer US.

Chi-Square Calculator by Social Science Statistics (2017). http://www.socscistatistics.com/tests/chisquare/Default.aspx

Christensen, E.E., Murry, R.C., Holand, K., Reynolds, J., Landay, M.J., & Moore, J.G. (1981). The effect of search time on perception. *Radiology, 138*, 361-365.

Craik, F., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671-684.

Crowley, R. S., Naus, G. J., Stewart, J., & Friedman, C. P. (2003). Development of Visual Diagnostic Expertise in Pathology: An Information-processing Study. *Journal of the American Medical Informatics Association*, *10*(1), 39–51.

Custers, E. J. F. M. (2008). Long-term retention of basic science knowledge: a review study. *Advances in Health Sciences Education*, *15*(1), 109–128.

de Groot, A. D. (1965). *Thought and Choice in Chess*. The Hauge: Mouton Publishers.

Dunphy, B. C., & Williamson, S. L. (2004). In pursuit of expertise. Toward an educational model for expertise development. *Advances in Health Sciences Education : Theory and Practice, 9*(2), 107–127.

Ericsson, K. A. (2004). Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Academic Medicine*, *79*(10 Suppl), S70–81.

Ericsson, K. A. (2006). Protocol analysis and expert thought: Concurrent verbalizations of thinking during experts' performance on representative tasks. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge Handbook of Expertise and Expert Performance* (pp. 223–241). New York: Cambridge University Press.

Ericsson, K. A. (2007). An expert-performance perspective of research on medical expertise: the study of clinical performance. *Medical Education*, *41*(12), 1124–1130.

Ericsson, K. A. (2008). Deliberate Practice and Acquisition of Expert Performance: A General Overview. *Academic Emergency Medicine*, *15*(11), 988–994.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*(3), 215–251.

Evans, K. K., Cohen, M. A., Tambouret, R., Horowitz, T., Kreindel, E., & Wolfe, J. M. (2011). Does visual expertise improve visual recognition memory? *Attention, Perception, & Psychophysics, 73*(1), 30–35.

Flin, R., Patey, R., Jackson, J., Mearns, K., & Dissanayaka, U. (2009). Year 1 medical undergraduates' knowledge of and attitudes to medical error. *Medical Education, 43*(12), 1147–1155.

Fonteyn, M. E., Kuipers, B., & Grobe, S. J. (1993). A description of think aloud method and protocol analysis. *Qualitative Health Research*, *3*(4), 430–441.

Gobet, F., & Simon, H. A. (1996). Recall of random and distorted chess positions: implications for the theory of expertise. *Memory & Cognition*, *24*(4), 493–503.

Haist, F., Shimamura, A. P., & Squire, L. R. (1992). On the relationship between recall and recognition memory. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18*(4), 691–702.

Helfer, R. E. (1972). Peer evaluation: its potential usefulness in medical education. *British Journal of Medical Education, 6*(3), 224–231.

Helms, C.A. (1995). Malignant bone tumors. In *Fundamental of Skeletal Radiology, Second Edition* (pp. 34-55). Philadelphia: W.B. Saunders Company.

Hmelo-Silver, C. E., Nagarajan, A., & Day, R. S. (2002). "It's harder than we thought it would be": A comparative case study of expert-novice experimentation strategies. *Science Education*, *86*(2), 219–243.

Johnson, E.J. (1988) Expertise and Decision Under Uncertainty: Performance and Process. In Chi, M. T., Glaser, R., & Farr, M. J. (Eds.) *The nature of expertise* (pp. 209-228). Hillsdale, NJ: Lawrence Erlbaum Associates.

Jonassen, D. H. (2000). Toward a design theory of problem solving. *Educational technology research and development*, *48*(4), 63-85.

Joorabchi, B., & Chawhan, A. R. (1975). Multiple choice questions. The debate goes on. *British Journal of Medical Education*, *9*(4), 275–280.

Kintsch, W. (1970). Models for free recall and recognition. In D. A. Norman (Ed.), *Models of Human Memory* (pp. 331–373). New York: Academic Press.

Kundel, H. L. (2006). History of Research in Medical Image Perception. *Journal of the American College of Radiology*, *3*(6), 402–408.

Kundel, H. L., Nodine, C. F., Conant, E. F., & Weinstein, S. P. (2007). Holistic Component of Image Perception in Mammogram Interpretation: Gaze-tracking Study. *Radiology, 242*(2), 396–402.

Laerd Statistics (2015). Statistical tutorials and software guides. Retrieved from https://statistics.laerd.com/

Lesgold, A. M. (1983). Acquiring Expertise (pp. 1–51). Office of Naval Research.

Lesgold, A., Rubinson, H., Feltovich, P., Glaser, R., Klopfer, D., & Wang, Y. (1988). Expertise in a Complex Skill: Diagnosing X-Ray Pictures. In Chi, M. T., Glaser, R., & Farr, M. J. (Eds.) *The nature of expertise* (pp. 311–342). Hillsdale, NJ: Lawrence Erlbaum Associates.

Linn, B. S., Arostegui, M., & Zeppa, R. (1976). Peer and self assessment in undergraduate surgery. *Journal of Surgical Research, 21*(6), 453–456.

Manning, D. (2010). Cognitive factors in reading medical images: A survey of cognitive factors and models of medical image interpretation. In E. Samei & E. Krupinski (Eds.), *The handobook of medical image perception and techniques* (pp. 91–106). New York: Cambridge University Press.

Mylopoulos, M., & Regehr, G. (2009). How student models of expertise and innovation impact the development of adaptive expertise in medicine. *Medical Education*, *43*(2), 127–132.

Nodine, C. F., Mello-Thoms, C., Kundel, H. L., & Weinstein, S. P. (2002). Time course of perception and decision making during mammographic interpretation. AJR Am J Roentgenol, 179(4), 917–923.

Norcini, J. J. (2003). Peer assessment of competence. *Medical Education, 37*, 539-543.

Raaijmakers, J. G., & Shiffrin, R. M. (1992). Models for recall and recognition. *Annual Review of Psychology, 43*, 205–234.

Rikers, R. M. J. P., Schmidt, H. G., & Boshuizen, H. P. A. (2000). Knowledge Encapsulation and the Intermediate Effect. *Contemporary Educational Psychology*, *25*(2), 150–166.

Salkowski, L.R. & Russ, R. (2017). Expert and novice differences in correlating anatomy and imaging. *In process.*

Schmidt, H. G., & Boshuizen, H. P. (1993). On the origin of intermediate effects in clinical case recall. *Memory & Cognition*, *21*(3), 338–351.

Schmidt, H. G., Norman, G. R., & Boshuizen, H. P. (1990). A cognitive perspective on medical expertise: theory and implication. *Academic Medicine*, *65*(10), 611–621.

Schmidt, H. G., & Rikers, R. M. J. P. (2007). How expertise develops in medicine: knowledge encapsulation and illness script formation. *Medical Education*, *41*(0), 1133–1139.

Snowden, P. T., Davies, I., & Roling, P. (2000). Perceptual learning of the detection of features in X-ray images: a functional role for improvements in adults' visual sensitivity? *Journal of Experimental Psychology: Human Perception and Performance, 26* (1), 379-390.

Ten Cate, O., & Durning, S. (2007). Peer teaching in medical education: twelve reasons to move from theory to practice. *Medical Teacher, 29*(6), 591–599.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80* (5), 352-373.

Tversky, B. (1973). Encoding processes in recognition and recall. *Cognitive Psychology, 5*, 275-287.

[Untitled illustration of recognition and recall]. Retrieved April 11, 2017 from webdesign-review.blogspot.com

Vogt, S. & Magnussen, S. (2005). Hemispheric specialization ad recognition memory for abstract and realistic pictures: A comparison of painters and laymen. *Brain and Cognition, 58*(3), 324.

Young, J. Q., Van Merrienboer, J., Durning, S., & Cate, Ten, O. (2014). Cognitive Load Theory: Implications for medical education: AMEE Guide No. 86. *Medical Teacher*, *36*(5), 371–384.

Chapter 5 References

Albanese, M. (2000). Problem-based learning: why curricula are likely to show little effect on knowledge and clinical skills. *Medical Education*, *34*(9), 729–738.

Colliver, J. A. (2000). Effectiveness of problem-based learning curricula: research and theory. *Academic Medicine*, *75*(3), 259–266.

Epstein, R. (2004). Learning from the problems of problem-based learning. *BMC Medical Education*, *4*(1), 1.

Nandi, P. L., Chan, J. N., Chan, C. P., Chan, P., & Chan, L. P. (2000). Undergraduate medical education: comparison of problem-based learning and conventional teaching. *Hong Kong Medical Journal*, *6*(3), 301–306.