

EVALUATING THE APPROPRIATENESS OF ELECTRONIC INFORMATION  
RESOURCES FOR LEARNING IN PRE-CLERKSHIP MEDICAL EDUCATION:  
AN EYE TRACKING STUDY

---

A Dissertation  
presented to  
the Faculty of the Graduate School  
University of Missouri-Columbia

---

In Partial Fulfillment  
of the Requirements for the Degree  
Doctor of Philosophy

---

by  
DINARA SAPAROVA  
Dissertation Supervisor – Dr. Joi L. Moore

DECEMBER 2015

The undersigned, appointed by the dean of the Graduate School, have examined the dissertation entitled

EVALUATING THE APPROPRIATENESS OF ELECTRONIC INFORMATION  
RESOURCES FOR LEARNING IN PRE-CLERKSHIP MEDICAL EDUCATION:  
AN EYE TRACKING STUDY

presented by Dinara Saparova,

a candidate for the degree of Doctor of Philosophy,

and hereby certify that, in their opinion, it is worthy of acceptance.

---

Dr. Joi L. Moore (Chair)

---

Dr. Sanda Erdelez

---

Dr. Gail Fitzgerald

---

Dr. Suzanne Boren

*To those  
who have supported me  
during this journey....*

## **ACKNOWLEDGEMENTS**

I am very lucky to have some of the brightest people on my dissertation committee who have been providing valuable guidance and advice during the years of the program and especially during the writing of this dissertation. I also could not have made it without the tremendous support of my family, colleagues, and friends. My deepest gratitude goes to my husband, my parents, and my University of Missouri colleagues. Special thanks to Nathaniel Nolan, Neeley Current, Lada Micheas, Dr. Kimberly Hoffman, Dr. Michael Hosokawa, and Dr. Jeffrey Belden; all of my University of Missouri friends; and finally to little Mathew who has shared with me nine months of the dissertation writing journey and daily office time.

## TABLE OF CONTENTS

ACKNOWLEDGEMENTS .....	ii
LIST OF ILLUSTRATIONS .....	v
LIST OF TABLES .....	vii
ABSTRACT .....	viii
<b>CHAPTER 1 – INTRODUCTION .....</b>	<b>1</b>
Problem Statement .....	1
Purpose of the Study .....	10
Research Questions .....	11
Significance of the Study .....	14
Outline of Research Methods .....	15
Assumptions, Delimitations, and Limitations of the Study .....	16
<b>CHAPTER 2 – LITERATURE REVIEW .....</b>	<b>18</b>
Theoretical Perspectives on Learning in Hypermedia .....	18
Principles of Cognitive Architecture and Cognitive Load Theory .....	21
Cognitive Load in Hypermedia Learning .....	22
The Cognitive Theory of Multimedia Learning in Hypermedia Learning .....	32
Measurements of Cognitive Load .....	35
<b>CHAPTER 3 – METHODS .....</b>	<b>45</b>
Research Objective and Questions .....	45
Study Setting and Participants .....	45
Materials and Instruments .....	46
Apparatus .....	52
Procedure .....	55
Data Analysis .....	62
<b>CHAPTER 4 – RESULTS .....</b>	<b>74</b>
Participants’ Demographics .....	74
Survey Results .....	74
Research Question 1 .....	78

Research Question 2 .....	84
Research Question 3 .....	89
<b>CHAPTER 5 – DISCUSSION .....</b>	<b>114</b>
Research Question 1 .....	114
Research Question 2 .....	119
Research Question 3 .....	121
Limitations, Implications, and Future Research .....	131
Conclusions.....	135
<b>APPENDICES .....</b>	<b>137</b>
Appendix A: Demographic Information Survey.....	137
Appendix B: Preferences for Electronic Information Resources Survey .....	139
Appendix C: Screening Test .....	140
Appendix D: Knowledge Tests.....	141
Appendix E: Interview Protocol .....	154
<b>REFERENCES .....</b>	<b>156</b>
<b>VITA .....</b>	<b>170</b>

## LIST OF ILLUSTRATIONS

FIGURE 1. NESTING OF NEW MEDIA .....	19
FIGURE 2. TWO-STEP QUESTION ORGANIZATION .....	48
FIGURE 3. TOBII X2-60 EYE TRACKER.....	53
FIGURE 4. CORNEAL REFLECTION AND BRIGHT PUPIL.....	54
FIGURE 5. TWELVE POSSIBLE REPLICATIONS OF 3 X 3 LATIN SQUARES .....	58
FIGURE 6. PERMUTATIONS OF THE THREE DISEASES .....	58
FIGURE 7. OVERVIEW OF THE EXPERIMENTAL PROCEDURE .....	61
FIGURE 8. AOIs IN <i>UPTODATE</i> .....	64
FIGURE 9. AOIs IN <i>ACCESS MEDICINE</i> .....	65
FIGURE 10. AOIs IN <i>WIKIPEDIA</i> .....	66
FIGURE 11. STUDENTS' REPORTED PREFERENCES FOR ELECTRONIC INFORMATION RESOURCES .....	75
FIGURE 12. STUDENTS' REPORTED CONFIDENCE IN THE APPROPRIATENESS OF ELECTRONIC INFORMATION RESOURCES FOR LEARNING .....	78
FIGURE 13. PUPIL DIAMETER FLUCTUATIONS AMONG PARTICIPANTS USING <i>UPTODATE</i> ....	81
FIGURE 14. PUPIL DIAMETER FLUCTUATIONS AMONG PARTICIPANTS USING <i>ACCESS MEDICINE</i> .....	82
FIGURE 15. PUPIL DIAMETER FLUCTUATIONS AMONG PARTICIPANTS USING <i>WIKIPEDIA</i> ....	82
FIGURE 16. EXAMPLE OF A LONGER FIXATION DURATION .....	83
FIGURE 17. DISTRIBUTION OF KNOWLEDGE TEST SCORES .....	85
FIGURE 18. GROUP 2 PREFERENCES FOR ELECTRONIC INFORMATION RESOURCES FOR LEARNING .....	88

FIGURE 19. TYPES OF HYPERLINKS IN <i>UpToDate</i> .....	90
FIGURE 20. EXAMPLE OF A QUICK READ FROM <i>WIKIPEDIA</i> .....	92
FIGURE 21. EXAMPLE OF VIEWING IMAGE IN <i>UpToDate</i> .....	94
FIGURE 22. LAYOUT OF MEDIA IN <i>ACCESS MEDICINE</i> .....	95
FIGURE 23. LAYOUT OF MEDIA IN <i>WIKIPEDIA</i> .....	96
FIGURE 24. SEARCH PAGE IN <i>UpToDate</i> .....	97
FIGURE 25. SEARCH PAGE IN <i>ACCESS MEDICINE</i> .....	98
FIGURE 26. EXAMPLE OF VIEWING SEARCH PAGE IN <i>UpToDate</i> .....	99
FIGURE 27. EXAMPLE OF VIEWING SEARCH PAGE IN <i>ACCESS MEDICINE</i> .....	100
FIGURE 28. LEFT-SIDE NAVIGATION MENU IN <i>UpToDate</i> .....	101
FIGURE 29. EXAMPLE OF STUDENT INTERACTIONS WITH NAVIGATION IN <i>ACCESS MEDICINE</i> .....	102
FIGURE 30. EXAMPLE OF STUDENT INTERACTIONS WITH NAVIGATION IN <i>WIKIPEDIA</i> .....	103
FIGURE 31. EXAMPLE OF READING THE SUMMARY IN <i>UpToDate</i> .....	108
FIGURE 32. DROP-DOWN SECTIONS IN <i>ACCESS MEDICINE</i> .....	111

## LIST OF TABLES

TABLE 1. EXAMPLES OF RESEARCH ON THE APPROPRIATENESS OF UPToDate, ACCESS MEDICINE, AND WIKIPEDIA .....	4
TABLE 2. INITIAL REVIEWER AGREEMENT ON THE KNOWLEDGE TEST ANSWERS .....	50
TABLE 3. PILOT TEST FEEDBACK .....	51
TABLE 4. EXAMPLE OF 3 X 3 LATIN SQUARE.....	57
TABLE 5. SIX LATIN SQUARES USED IN THE EXPERIMENT .....	59
TABLE 6. THREE LATIN SQUARES USED WITH GROUP 2 .....	71
TABLE 7. REPORTED PURPOSE, FREQUENCY, AND REASONS FOR USING ELECTRONIC INFORMATION RESOURCES .....	75
TABLE 8. NORMALITY TEST OF EYE METRICS ASSOCIATED WITH MENTAL EFFORT .....	79
TABLE 9. SIGNIFICANCE OF EYE METRICS ASSOCIATED WITH MENTAL EFFORT .....	80
TABLE 10. NORMALITY TEST OF THE KNOWLEDGE TEST VARIABLE .....	84
TABLE 11. REPORTED LEVELS OF FAMILIARITY AND PRIOR KNOWLEDGE OF THE DISEASES .....	86
TABLE 12. AVERAGE NUMBER OF VISITED UNIQUE PAGES .....	87

EVALUATING THE APPROPRIATENESS OF ELECTRONIC INFORMATION  
RESOURCES FOR LEARNING IN PRE-CLERKSHIP MEDICAL EDUCATION:  
AN EYE TRACKING STUDY

Dinara Saparova

Dissertation Supervisor – Dr. Joi L. Moore

**ABSTRACT**

Current U.S. medical students have begun to rely on electronic information repositories – such as *UpToDate*, *Access Medicine*, and *Wikipedia* – for their pre-clerkship medical education. However, it is unclear whether these resources are appropriate for this level of learning due to factors involving information quality, level of evidence, and the requisite knowledge base. This study evaluated the appropriateness of electronic information resources from a novel perspective: the amount of mental effort learners invest in interactions with these resources and the effects of the experienced mental effort on learning.

Eighteen first-year medical students read about three unstudied diseases in the three above-mentioned resources (a total of 54 observations). Their eye movement characteristics (i.e., fixation duration, fixation count, visit duration, and task-evoked pupillary response) were recorded and used as psychophysiological indicators of the experienced mental effort. Post reading, students' learning was assessed with a multiple-choice test. Eye metrics and test results constituted quantitative data that were analyzed according to the repeated Latin square design. Students' perceptions and observations of

their interactions with the information resources constituted qualitative data that were also obtained. Participants' feedback from semi-structured interviews and recordings of students' information acquisition behaviors were reviewed, transcribed, and open coded for the emergent themes.

Compared to *Access Medicine* and *Wikipedia*, *UpToDate* was associated with significantly higher values of eye metrics suggesting higher mental effort experienced by learners when using this resource. No statistically significant difference between the amount of mental effort and learning outcomes was found. More so, descriptive statistical analysis of the knowledge test scores suggested similar level of learning regardless of the information resource used. Students' feedback and observations of their behaviors were informative in understanding and interpreting the differences in quantitative findings.

Judging by the learning outcomes, all three information resources were found appropriate for learning. *UpToDate*, however, when used alone, may be less appropriate for first-year medical students' learning as it does not fully address their information needs and is more demanding in terms of invested cognitive resources.

## **CHAPTER 1 – INTRODUCTION**

### **Problem Statement**

#### **Problem-Based Learning and Information Literacy**

In many medical schools in the United States, contemporary medical education during pre-clerkship years, after undergoing significant changes, now prioritizes a learner-centered approach called problem-based learning (PBL) (Barrows, 2000). PBL highlights learning about a subject by solving real life problems and is intended to inspire students to ask questions, integrate theory and practice, and demonstrate the application of knowledge and skills in search of feasible solutions to the learning issues (Savery, 2006).

The PBL approach in medicine also admits that no person can retain all the information that constitutes medicine today and emphasizes the importance of the skills that help learners find information necessary to address their information needs (Woods, 1994) and accentuates the importance of students' self-directed and self-regulated learning (Blumberg & Michael, 1992). As a result, medical students undergoing PBL are expected to be discerning information seekers with information seeking behaviors and literacy standards favorably distinguishing them from traditional medical curriculum students (Dodd, 2007; Eskola, 2005).

A number of studies investigated the effect of PBL on information-seeking behavior of medical students and identified the following patterns:

- Medical students in PBL use a greater number and variety of resources (Cooper & Elnicki, 2011; Dodd, 2007; Deretchin, Yeoman, & Seidel, 1998; Lawrence, Grosenick, Simpson, & Van Susteren, 1992).

- Medical students in PBL tend to self-select information resources (Dodd, 2007; Rankin, 1992).
- Medical students in PBL give preference to electronic resources (Chen, Lin, Chang, & Sun, 2011; Boumarafi, 2010; Peterson, Rowat, Kreiter, & Mandel, 2004; Deretchin et al., 1998; Rankin, 1992).

When it comes to students' decisions with regards to the selection and use of information resources, they are often based on normative beliefs, i.e., organizational culture and influence of the superiors (Tao, 2008). Let's take *UpToDate* as an example. The current popularity of *UpToDate* in medical practice is overwhelming. It is considered one of the leading information resources included in the evidence-based collections of health sciences and medical libraries and is recommended as a reliable, trustworthy information resource for patient care (Guyatt, Meade, Rennie, & Cook, 2008; Cooper & Elnicki, 2011; Peterson et al., 2004). The fact that *UpToDate* is being favored by many physicians and residents makes it 'a gold standard' among the US medical students as well (Gaitsgory, Burgess, & Mellis, 2013), which, however, raises concerns about its suitability for learning in pre-clerkship curriculum.

When students decide to use information resources for learning based on normative beliefs rather than clear understanding of their suitability, they may experience a mismatch between the intended and the actual use of the information resource. This mismatch reflects students' cognitive unpreparedness to successfully utilize the information resource for learning. To avoid the negative experience, it is, therefore, imperative for students and their mentors to be aware of the suitability of most commonly used information resources for learning during various phases of professional education

and able to differentiate between the costs and benefits of using certain information resources.

### **Electronic Information Resources Most Commonly Used in Medical Education**

In this dissertation I focus on evaluating the appropriateness of electronic information resources that are most commonly used by medical students in the U.S. medical schools for learning. Among them are electronic textbooks (i.e., *Access Medicine*) (Chen et al., 2011; Boumarafi, 2010; Peterson et al., 2004; Cogdill & Moore, 1997), textbook-like resources (i.e., *UpToDate*) (Cooper & Elnicki, 2011; Peterson et al., 2004), and electronic encyclopedia (e.g., *Wikipedia*) (Crawford, 2001; Head & Eisenberg, 2010; Hughes, Joshi, Lemonde, & Wareham, 2009; Lim, 2009).

Electronic textbooks (also referred to as e-books or multimedia textbooks) are the result of integrating classical book structure with features provided within an electronic environment such as interactivity and ability to read on a computer (Landoni, 2003). To provide faster and more extensive searching, e-books can be grouped into either collections of textbooks in digital formats (e.g., *Access Medicine*) or textbook-like resources (e.g., *UpToDate*) (Guyatt, Meade, Rennie, & Cook, 2008). The former may include traditional printed texts available in digital formats and resources published completely online (Heyd, 2010) while the latter typically include clinical evidence summarized and integrated with other types of information from lower evidence levels resources (Guyatt et al., 2008).

*Wikipedia* remains to be a popular resource among medical students, particularly during the first couple of years of program. Allahwala, Nadkarni, & Sebaratnam (2013) found a significant correlation between the year of medical school and the use of

*Wikipedia* as the first resource, with more senior students less likely using *Wikipedia* as initial resource. Judd & Kennedy (2011) demonstrated that *Wikipedia* was the second most used information resource after Google among medical students over a period of three years with the first two years using *Wikipedia* more often. It is likely that popularity of *Wikipedia* among students is explained by its features such as ease of access, scope of coverage, currency of information, and usable interface (Head & Eisenberg, 2010).

Previous research has provided evidence as to the appropriateness of the above-mentioned information resources for learning and medical practice from several perspectives (Table 1). This body of evidence, however, remains sporadic and somewhat controversial.

Table 1

*Examples of Research on the Appropriateness of UpToDate, Access Medicine, and Wikipedia*

<b>Information resource</b>	<b>Research perspective</b>
<i>UpToDate</i>	Information retrieval accuracy (Fenton & Badgett, 2007); information retrieval speed (Ahmadi et al., 2011); user satisfaction (Campbell & Ash, 2006); students' information needs (Peterson et al., 2004).
<i>Access Medicine</i>	Accuracy of information and coverage (Pender, Lassere, Kruesi, Del Mar, & Anuradha, 2008); duration of information episodes and individual preferences (Hayward, El-Hajj, Voth, & Deis, 2006); clinical information needs (Esparza, Shi, McLarty, Comegys, & Banks, 2013).
<i>Wikipedia</i>	Accuracy, completeness, and quality of information (Lavsa, Corman, Culley, & Pummer, 2011; Andrew Wood & Struthers, 2010); depth of

---

	discussions of mechanisms and pathogenesis of diseases (Azer, 2014); quality of references (Haigh, 2011; Lavsa et al., 2011); level of evidence (Pender et al., 2008).
--	--

---

## **Electronic Information Resources as Examples of Hypermedia**

In this study, I propose to consider electronic information resources most commonly used in medical education (i.e., *UpToDate*, *Access Medicine*, and *Wikipedia*) as examples of hypermedia. This proposition is rooted in the fact that features and functionality of these information resources concur with the features and functionality of hypermedia learning environments.

Conklin (1987) defined hypermedia as network-like information structures, in which pieces of information are stored in interconnected nodes and can be accessed via hyperlinks. Functional organization of the identified electronic information resources clearly shows presence of hyperlinks that connect information pieces stored in separate nodes. Another distinguishing feature of hypermedia environments is availability of multimedia elements such as audio recordings, graphics, and animation (Conklin, 1987). Overview of interfaces of the three resources demonstrates presence of images, animations, and audio recordings. Finally, in hypermedia learning environments the control over the order, selection, and presentation of information is established by the learner rather than the system (Scheiter & Gerjets, 2007). When it comes to learner interaction with the above-mentioned electronic resources, learner control is apparent, as it is the learner who makes navigational decisions and decisions about the order of information to be accessed and used.

Due to hypermedia characteristics, learning inside these environments has its own advantages and disadvantages. Among the definite pluses is the opportunity for learners to construct their own paths to understanding through the extensive use of hyperlinks and multiple formats of information presentation, which is known to foster cognitive flexibility (Jacobson & Spiro, 1995). Further, by engaging in knowledge construction through integration of new and existing information, learners benefit from constructivist rather than objectivist learning (Jonassen, 1999). Finally, by following non-linear information structures, learners are encouraged to establish relationships between the concepts, which also supports active knowledge construction (Kintsch, 1998).

On the other hand, the same structural peculiarities of hypermedia and emphasis on learner control may impede learning. For example, Conklin (1987) explained that navigation planning can occupy cognitive resources of the working memory, making them unavailable for efficient information processing, and referred to this phenomenon as cognitive overhead. This is especially true for learners with low prior knowledge, as they are more likely to experience difficulties with the selection of information due to being unable to predict their information needs and understand relationships between the concepts (Norman, 1982). Lack of situational models may make such learners follow hyperlinks purely based on their interests, which, on one hand, may result in discovery and incidental learning (Heller, 1990) but, on the other hand, cause interruptions in information processing (Lorch, 1993) and orientation problems known as the phenomenon of being lost-in-hyperspace (Conklin, 1982).

## **Cognitive Load in Hypermedia**

One of the leading theories that found a wide application to hypermedia learning environments is the theory of cognitive load (DeStefano & LeFevre, 2007; Zumbach & Mohraz, 2008). The theory operates under the assumption of the limited processing capacity of the working memory. This means that for learning to be successful, cognitive resources should primarily be expended on the processes contributing to knowledge construction. Central to cognitive load theory (CLT) is a multidimensional construct of cognitive load, which signifies different types of learner's cognitive resources needed to accommodate the demands of the task.

In electronic learning environments, including hypermedia, performance on the tasks can be associated with three types of cognitive load (Sweller, Merrienboer, & Paas, 1998). First, cognitive resources needed to accommodate for the inherent characteristics of the learning material, known as intrinsic cognitive load. Second, cognitive resources needed to establish the relationship between existing and new knowledge and comprehend the material, known as germane cognitive load. Third, cognitive resources needed to access information, known as extraneous cognitive load.

Specifically to hypermedia, it is likely that both intrinsic and germane cognitive loads will be high if learning extends beyond merely establishing the relationships between the concepts (Zumbach, 2006; Antonenko & Niederhauser, 2010). Extraneous cognitive load can be high as well if learners spend their cognitive resources on deciphering meaning during reading (Antonenko & Niederhauser, 2010); understanding and interpreting information organization (Geissler, Zinkhan, & Watson, 2001); and overcoming the barriers associated with the ease of navigating and operating the system (Niederhauser, 2007).

The three types of cognitive load are considered additive (van Merriënboer & Sweller, 2010), i.e., reduction or increment in one type is likely to cause fluctuations in the other two. This means that whether the amount of cognitive resources required for overcoming extraneous cognitive load will be detrimental for learning depends on the amount of cognitive resources required for overcoming intrinsic or germane cognitive load. If learning requires a significant amount of cognitive resources for understanding difficult material, cognitive resources expended on overcoming the barriers associated with access to this material should be reduced to a minimum.

From the three types of cognitive load, it is extraneous cognitive load that is considered to be the most interfering with efficient information acquisition and processing and potentially the most detrimental for learning. Extraneous cognitive load is also considered to be the easiest to manipulate through instructional design (Paas et al., 2003). Effective manipulations can reduce extraneous cognitive load, free capacity for germane cognitive load, and improve learning by reducing overall cognitive load.

### **Eye Tracking as a Technique for Measuring Cognitive Load**

There are a variety of modalities allowing the measure of a learner's cognitive load associated with task performance, including the assessment of behavior, physiology, psychology, and learning outcomes (Brünken, Plass, & Leutner, 2010). Another widely used method for measuring cognitive load during task performance, which represents a combination of physiology and psychology, is eye tracking (Paas, Ayres, & Pachman, 2008). Eye tracking technology is used for tracking and recording eye movements of individuals in order to provide an inside look into people's visual processes (Duchowski, 2003). Visual processes, in their turn, are considered to be an excellent behavioral

measure of cognitive processes, including reading (Liversedge & Findlay, 2000; Rayner, 1998).

Two basic components of eye movement are fixations and saccades (Rayner, 2009). Fixations allow for new information to be acquired because the eye remains still; saccades allow for rapid eye movement, during which information acquisition is, typically, impossible. Among many characteristics of eye movements that are helpful in interpreting cognitive processes are fixation duration, fixation count, saccade length and direction (e.g., forward vs regressions), blink rate, and pupil response (Poole & Ball, 2005). For example, cognitive load imposed either by the information processing demands (e.g., germane cognitive load) or the difficulty of the material (e.g., intrinsic cognitive load) finds reflection in higher fixation frequency, especially when reading involves longer and less familiar words (Rayner, Sereno, & Raney, 1996). Higher rate of regressive saccades is usually associated with reading that involves dealing with syntactically ambiguous sentences (Paterson, Liversedge, & Underwood, 1999), which also increases information processing load. Low blink rate (Bruneau, Sasse, & McCarthy, 2002) and larger degree of pupil dilation (Iqbal, Zheng, & Bailey, 2004) are the indicators of high workload associated with rigorous information processing.

The idea of applying eye tracking for evaluating the effects of certain hypermedia features on the efficiency of information acquisition and processing has been explored previously. Hyönä & Lorch (2004) investigated the effect of topic headings in electronic texts on text processing efficiency; Leuthold, Schmutz, Bargas-Avila, Tuch, & Opwis (2011) explored the differences in menu design on user performance and subjective preferences; Holsanova, Holmberg, & Holmqvist (2008) studied the effects of reading

information graphics on spatial contiguity; and Goldberg (2012) examined visual organization of the content on perceived page complexity. Despite the availability of research evidence, studies utilizing eye-tracking to exploring the effects of hypermedia on the efficiency of information acquisition and processing remain in short demand. The primary reason for that is significant time- and labor-intensity associated with implementation of eye tracking methodology (Hyönä, 2010; van Gog & Scheiter, 2010).

### Purpose of the Study

This study is a contribution to the knowledge of the appropriateness of electronic information resources most commonly used in medical education (e.g., *UpToDate*, *Access Medicine*, and *Wikipedia*) from a novel perspective. Specifically, I propose to use the amount of learners' mental effort invested in information acquisition and processing from these electronic information resources as a measure of resource appropriateness.

Mental effort is a measurable dimension of cognitive load (along with mental load and performance) that represents the cognitive capacity allocated by the learner when working on the task (Paas & van Merriënboer, 1994). It is believed to represent the actual cognitive load experienced by the learner and yields the most reliable estimate of cognitive load<sup>1</sup> (Paas, Ayres, & Pachman, 2008).

Even a quick look at the interfaces of the three information resources discussed above reveals differences in their information organization and presentation. For example, *Wikipedia* contains several types of immediately available multimedia (e.g., images, animations, audio recordings) and hyperlinks embedded in each page. *UpToDate* offers more condensed content, primarily textual, and almost no multimedia immediately

---

<sup>1</sup> Throughout the study I use the terms 'mental effort' and 'cognitive load' interchangeably.

accessible from the pages. *Access Medicine*'s distinctive feature is the presentation of content (both text and images) in the center on the page with large white margins on both sides. It also has a minimum number of hyperlinks.

The differences in content organization and presentation among the three information resources help form the main assumption of the study: *UpToDate*, *Access Medicine* and *Wikipedia* are likely to differ in terms of extraneous cognitive load they impose on the learner. Expected differences in extraneous cognitive load, in their turn, are likely to result in differences in mental effort required from learners to access, acquire, and efficiently process information from each resource. Further, according to the principle of additivity of cognitive load, the amount of mental effort expended on overcoming sources of extraneous cognitive load could interfere with the amount of mental effort expended on knowledge construction and, ultimately, learning, especially if learning requires establishing the relationships between the concepts. Therefore, it is expected that learning from information resources associated with higher extraneous cognitive load is likely to be inferior. Both the amount of mental effort invested in interactions with information resources and the resulting quality of learning will serve as the bases for making conclusions about the appropriateness of the three electronic information resources for learning.

### **Research Questions**

Based on the described assumptions, the following research questions were formed:

1. What are the differences in learners' mental effort invested in interactions with electronic information resources?

2. How do the differences in learners' mental effort invested in interactions with electronic information resources affect learning?
3. How do certain interface design characteristics of electronic information resources contribute to or interfere with information acquisition and, consequently, learning?

### **Theoretical Framework**

The theory chosen to guide this research is the cognitive load theory (CLT) (Sweller, 1988). The theory asserts that learning is obstructed when the capacity of the working memory is exceeded. CLT distinguishes between the three types of cognitive load: intrinsic load – associated with overcoming the inherent complexity of the learning material; extraneous cognitive load – associated with overcoming material presentation; and germane cognitive load – associated with the effort spent on information processing and construction of meaning. Because the types of cognitive load are considered highly additive (van Merriënboer & Sweller, 2010), it is recommended to keep extraneous cognitive load at a minimum, especially when intrinsic cognitive load is expected to be high, in order to free the capacity of the working memory and increase germane cognitive load.

Despite the existing differentiation between the three types of cognitive load in nomenclature, in reality it is not always possible to clearly distinguish between them. In his seminal article, de Jong (2010) elaborated on three of such distinctions. When elaborating about the distinction between intrinsic and germane cognitive load, he wrote: “The first question is how intrinsic load, defined as a characteristic of the material, can contribute to cognitive load of the learner. A contribution to cognitive load cannot come

from the material as such, but can only take place when the learner starts processing the material” (p. 110). This means that cognitive load begins to exist when the learner begins to process information, and that without any action on the learner part, there cannot be cognitive load.

With regard to the distinction between extraneous and germane cognitive load, de Jong built his argument by referring to several research studies. He exemplified that the difference between extraneous and germane load could be problematic because it seemed to depend on the expertise level of the learner and cited Paas, Renkl, & Sweller (2004) who wrote: “A cognitive load that is germane for a novice may be extraneous for an expert” (pp. 2-3). Among other studies he cited were Gerjets & Scheiter (2003) who proposed to consider germane processes as extraneous depending on the learning goal.

Finally, de Jong points out that the distinction between intrinsic and extraneous cognitive load could be problematic, too. This primarily comes from the fact that intrinsic load, being associated with the inherent difficulty of the material, cannot be changed. This conception, however, is questioned by the findings from the DeLeeuw and Mayer (2008)’s study who were able to manipulate intrinsic cognitive load by changing the complexity of the sentences (i.e., extraneous cognitive load) in a text while preserving the same content.

Difficulties in clear differentiation between the three types of cognitive load are likely to interfere with the interpretation of the experiments and the accuracy of cognitive load measurements. Because it is hard to accurately measure each particular type of cognitive load, conclusions about which of the three types of cognitive load is mostly contributing to the overall cognitive load are complicated as well. It is, however, believed

that with the introduction of new technology such as eye tracking and brain scans, it would become possible to observe patterns suggesting fluctuations in each of the three types of cognitive load and obtaining their accurate measurements (Jaeggi et al., 2007; Paas et al., 2008).

### **Significance of the Study**

The study is multidisciplinary in nature; it carries both theoretical and practical implications for several domains. Theoretically, this research makes a contribution to the existing pool of studies on information-seeking behavior of medical students in the context of problem-based learning. Specifically, it adds knowledge about behavioral characteristics of medical students during the use of certain electronic information resources for purposeful information acquisition. The study also contributes to the literature on hypermedia learning by suggesting to consider electronic information resources as examples of hypermedia learning environments. Further, the research expands application of cognitive load theory, instructional design principles, and usability heuristics to the context of electronic information resources, which has not been done previously. Finally, current investigation contributes to the knowledge of the effectiveness of the widely accepted and used psychophysiological measures of cognitive processes – eye tracking – in the novel contexts of medical education and electronic information resources.

The practical value of the research is in informing the parties involved in acquisition and provision of pre-clerkship curricula about the appropriateness of electronic information resources commonly used for learning. The findings could be helpful in suggesting a more critical approach to selection and use of available

information resources by weighing their potential costs and benefits. The knowledge from the study can also be helpful to the faculty, health science librarians, and medical school administration in guiding their decisions regarding the content of information literacy orientation for newly accepted medical students and decisions on purchasing online collections of information resources. Further, results of the evaluation of interfaces and functionality of information resources could be of value to the developers of these resources as they include suggestions for improvement and re-design.

## **Outline of Research Methods**

### **The Study Context**

The study was conducted during the summer semester of 2013-2014 academic year with 18 first-year medical students representing Class 2017 of the University of Missouri School of Medicine. By the time of the experiment, these students had completed the first year of the two-year pre-clerkship problem-based learning curriculum but had not yet started their second year of program. Due to the nature of the study, students who were experiencing any vision problems (e.g., complete or partial vision loss and prescription glasses of 4+ diopters) were not eligible for participation.

### **Research Methodology**

In this research both qualitative and quantitative methods of data collection and analyses were employed. Quantitative data included eye metrics such as total fixation duration, fixation count, total visit duration, and task-evoked pupillary response that were collected with the help of Tobii eye tracker. Repeated Latin square design (Natrella, 2010) was the primary method of quantitative data analysis, which was conducted using IBM SPSS Statistics 19 software application.

Qualitative data were collected during end-of-session, semi-structured interviews for the purpose of helping to interpret quantitative findings. Additional qualitative data resulted from observations of participants' information acquisition behavior during information resource use and evaluation of information organization and presentation in accordance with the principles of instructional design, multimedia learning, and usability heuristics. Qualitative data were reviewed several times for the occurrence of common themes and open coded in QDA Miner qualitative software.

### **Assumptions, Delimitations, and Limitations of the Study**

#### **Assumptions**

The study was conducted under the following assumptions.

1. Because the students who participated in this study successfully completed the first year of program, it was assumed that the average level of medical knowledge they acquired as part of the first year curriculum was approximately the same.
2. Because in the pre-test survey participants reported familiarity with all three information resources (i.e., *UpToDate*, *Access Medicine*, and *Wikipedia*), it was assumed that students were equally familiar with the features and functionality of these resources.

#### **Delimitations**

The study was limited in scope by the following delimitations:

1. No individual characteristics affecting reading behavior were taken into account because the sample size and recruitment method did not allow for equal variability.
2. Only three diseases were included in the experiment.

3. Students were limited to one information resource to learn about one disease at a time.
4. Students were given a limited time to learn about each disease (i.e., about 20-25 minutes).
5. Students were given the task of learning as much as they could about each of the diseases in the allocated amount of time.

## **Limitations**

The limitations of the study included the following:

1. The generalizability of the study was limited because the study focused only on students representing the first year of pre-clerkship curriculum from one academic institution. In addition, the sample of participants, even though sufficient for the type of the chosen analysis, could be viewed by some as relatively small. The sample was also generated based on convenience, which likely resulted in under- or over-representation of particular groups within the sample.
2. Experimental conditions could have affected the results of the study.
  - a. The timing of the study – the summer months – could have been an issue because some students were not in the ‘studying mode’, which could have affected their ability to remember or comprehend certain medical information.
  - b. To observe the validity of the eye tracking data, it was impossible to zoom in/out on the content available in information resources, which made it impossible for the participants to accommodate for the preferred font size.

## CHAPTER 2 – LITERATURE REVIEW

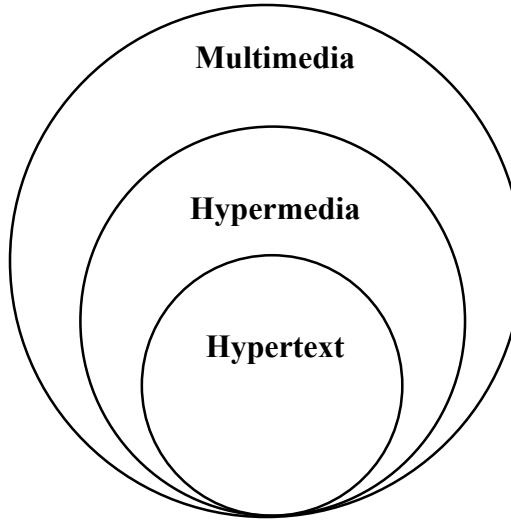
### Theoretical Perspectives on Learning in Hypermedia

Hypermedia learning environments can be defined as network-like information structures, in which pieces of information are stored in interconnected nodes and can be accessed by electronic hyperlinks (Conklin, 1987). The basic rationale behind hypermedia is that information can be presented in a non-linear format, which allows learners to make their way through these information-rich environments in their own, self-directed manner. If the system contains primarily textual information, it is referred to as hypertext (Gall & Hannafin, 1994); whereas if the system contains the elements that include “text, graphics, digitized speech, audio recordings, pictures, animation, film clips, and presumably tastes, odors, and tactile sensations, it is referred to as hypermedia” (Conklin, 1987, p. 18)<sup>2</sup>. Hypermedia systems, therefore, can be viewed as an extension of hypertext because they contain information in different formats (e.g., text, images, video, and audio). Due to presence of media elements and their non-linear connectivity, hypermedia can also be considered a special case of non-linear multimedia. The distinction between the two, however, is in the fact that in multimedia learning environments the control over the order, selection, and presentation of information is established by the system while in hypermedia environments it is established by the learner and is known as learner control (Scheiter & Gerjets, 2007) (Figure 1).

A few learning theories provide theoretical foundations with regards to learning in hypermedia. Cognitive Flexibility Theory suggests that hypermedia-based environments allow the learners to construct their own, unique pathways towards topic understanding

---

<sup>2</sup> Based on the interconnectedness between hypermedia and hypertext, I am using the terms ‘hypermedia’ and ‘hypertext’ interchangeably throughout my writing.



*Figure 1. Nesting of new media*  
 (adapted from Kommers, Grabinger, & Dunlap, 1996).

through the extensive use of available hyperlinks (Jacobson & Spiro, 1995). In other words, because the learners can access the same information several times and from several locations, they are likely to encounter this information from various perspectives and in various contexts. This process of ‘criss-crossing the conceptual landscape’ (Spiro, Coulson, Feltovich, & Anderson, 1988) promotes learner engagement in deeper cognitive processes by constantly comparing and contrasting the linked information. The criss-crossing of content, in its turn, facilitates deeper processing of information, as the learners continuously evaluate which information may be helpful in achieving their learning goals and how different information units are related.

By accessing new information, organizing it into coherent mental representations, and integrating it with existing knowledge, the learner gradually constructs new meanings and new knowledge. Such learner control allows the learners to proceed at their own pace and according to their own learning needs, which increases their interest in and motivation for learning (Snow, 1980). It is, therefore, believed that the unique way in

which hypermedia organizes information and makes it accessible for learners makes these environments highly conducive for constructivist rather than objectivist learning (Jonassen, 1999; Meyer, 1999).

The Construction-Integration Model (Kintsch, 1998) has been used to demonstrate how a purposeful use of hyperlinks may assist with active knowledge construction. According to this theory, text comprehension is achieved by sequentially forming a coherent mental representation from the text through cyclical processes of construction and integration. During the construction phase, the learner establishes a network of interconnected propositions from the text. During the integration phase, the learner connects ideas of the text and integrates them with the existing knowledge, thus, creating a situation model. It is, therefore, believed that by following non-linear information presentation pertinent to hypertext, learners are encouraged to establish inferences and relationships between different nodes and actively construct knowledge.

Recently, a number of authors have applied cognitive load theory to studying learning in hypermedia (DeStefano & LeFevre, 2007; Zumbach & Mohraz, 2008). Hypermedia environments allow for presentation of information in multiple formats that foster learning, e.g., visual and auditory, static and dynamic, and via hyperlinks that offer freedom of navigation. By providing learners with a variety of learning experiences, these environments allow for accommodation of individual learning needs, thus, facilitating knowledge construction and engagement in meaningful learning. At the same time, these environments are often designed in violation of the principles of human cognitive architecture, which carries the potential of causing the learner to experience cognitive overload and inefficient learning (Paas et al., 2008). Therefore, cognitive load theory is

one of the fundamental theories that can help explain how design of hypermedia environments affects learners' cognitive load and the quality of learning.

### **Principles of Cognitive Architecture and Cognitive Load Theory**

The principles of human cognitive architecture explain the roles of and relationships between working and long-term memory during learning (Sweller et al., 1998). Working memory is most commonly used to process information; hence, it is associated with cognitive activity, in which human beings engage consciously. Modern theories of working memory distinguish between visual working memory (e.g., images) and auditory working memory (e.g., text) that each processes information of a certain type separately (Baddeley, 1992). Further, because human beings are able to deal with only two or three items of information at the same time, the capacity of working memory is extremely limited (Miller, 1956).

Long-term memory, on the other hand, is associated with unconscious storage of information in the form of schemas. People become aware of the long-term memory content by filtering it through the working memory, when the relationships between new information and existing schemas are established (Chi, Glaser, & Rees, 1982). The more interrelated schemas a person has available in the long-term memory, the less load is imposed on the working memory, which allows for easier engagement in complex reasoning.

The limited processing capacity of working memory is the main premise upon which the cognitive load theory builds. According to this theory, different sources of cognitive load contribute to different types of cognitive load (Sweller et al., 1998). For example, inherent complexity of the material is determined by high element interactivity

between the constructs. In order to understand the material and process information efficiently, the learner needs to keep a large number of information units in the working memory. This cognitive load imposed by the complexity of the learning material is referred to as intrinsic cognitive load. The format in which information is organized and the manner in which it is presented may interfere with the efficient information processing in working memory. Because this type of cognitive load is often imposed by the presentation of the material rather than the material itself, it is referred to as extraneous cognitive load. Finally, learner's effort spent on processing information, constructing new meaning, and comprehending the material is referred to as germane cognitive load.

Sweller (1994) argued that the total cognitive load is a combination of intrinsic and extraneous cognitive loads and that both types of cognitive loads can be manipulated in the design and presentation of material. If the material is not appropriately designed, a combination of high intrinsic and extraneous cognitive load may overwhelm the limited capacity of the working memory. As a result, efforts spent on overcoming unnecessary cognitive load may result in decreased amounts of cognitive load necessary for constructing meaning and, consequently, impair learning.

### **Cognitive Load in Hypermedia Learning**

Hypermedia learning is likely to be associated with increased cognitive load for a number of reasons. First, learning with hypertext is likely to lead to an increased intrinsic cognitive load if the learning goes beyond just learning the facts and requires establishing the relationships between the concepts (Zumbach, 2006). Intrinsic cognitive load is likely to be higher if the interactivity between the information elements is high as well. The

higher element interactivity, the more complex the material is likely to be and the more difficult it is to understand.

The amount of intrinsic cognitive load experienced by the learners depends on their prior knowledge. As Spiro (1988) explained, when learners interact with hypertext, they have the opportunity to traverse the contextual landscape and achieve a deeper understanding of the material by accessing information from various perspectives and relating it to existing knowledge. The more conceptual schemas the learners have about a topic, the easier it is for them to see how these schemas are interrelated and how they contribute to the formation of a single element. These schemas, or prior knowledge, allow learners to organize information in meaningful ways and help integrate and organize existing and new information (Chi et al., 1982). Therefore, learners with high prior knowledge have the ability to combine multiple information units together and view them as single element, which makes learners with high prior knowledge less prone to intrinsic load when dealing with complex material (Gerjets, Scheiter, Opfermann, Hesse, & Eysink, 2009).

Learner's effort spent on processing information, constructing new meaning, and comprehending the material is referred to as germane cognitive load. Germane cognitive load is, therefore, the load placed on the working memory in order to deal with intrinsic cognitive load that leads to learning (van Merriënboer & Sweller, 2010). In hypermedia learning, increased cognitive flexibility associated with building a situation model is likely to contribute to the increase in germane cognitive load (Antonenko & Niederhauser, 2010). That is, if understanding the connections between the concepts is beyond learner's aptitude, the germane cognitive load is likely to increase and interfere

with learning by occupying a significant portion of working memory capacity. On the other hand, if learning is not challenging enough, germane cognitive load is likely to be low, resulting in a loss of interest in learning (Antonenko & Niederhauser, 2010).

### **Sources of Extraneous Cognitive Load**

Extraneous cognitive load is the type of cognitive load imposed by the manner and format of material presentation; therefore, it is an important part of instructional design science. Extraneous cognitive load results from unnecessary cognitive processes as a result of interacting with the material that do not lead to schema acquisition (Sweller et al., 1998). Cognitive load theory assumes that intrinsic and extraneous cognitive loads are additive (van Merriënboer & Sweller, 2010). This means that whether extraneous cognitive load is going to be a potential problem depends on the intrinsic cognitive load. For example, if intrinsic cognitive load is low, high extraneous cognitive load resulting from inadequate instructional design is not going to be harmful for learning because the total cognitive load stays within the working memory capacity. On the other hand, high intrinsic cognitive load in combination with high extraneous cognitive load may result in cognitive overload by exceeding working memory resources. The more extraneous cognitive load is reduced, the more working memory resources can be allocated to overcoming intrinsic cognitive load and inducing germane cognitive load necessary for learning (van Merriënboer & Sweller, 2010).

In hypermedia, extraneous cognitive load may exist on several levels. First, since learning from hypertext involves reading, extraneous cognitive load can be associated with basic reading processes such as letter recognition and decoding words, comprehension monitoring, and situational model development (Antonenko &

Niederhauser, 2010). Second, the amount of extraneous cognitive load depends on learner's familiarity with the system and ease of navigating and operating it (Niederhauser, 2007). Other sources of extraneous cognitive load are described below.

**Visual complexity.** Design of hypermedia learning environments and content organization play an important role in regulating the amounts of extraneous cognitive load imposed on the learners. Miniukovich & De Angeli (2014) proposed to consider information organization, information amount, and information discriminability as the main determinants of visual complexity. Among the most influential interface design elements contributing to perceived visual complexity Geissler et al. (2001) found to be the number of links and graphic elements, large text blocks, page length, and animations. Subjective user perceptions of visual complexity were also linked to information density, diversity, and position of the elements and proposed to be used as an implicit measure of cognitive load (Harper, Michailidou, & Stevens, 2009).

Wästlund, Norlander, & Archer (2008) examined how page layout affected mental workload. Two types of text presentation were compared: one where text of the document fit on a screen and was exactly the same size as the computer screen; the other where text was presented in a continuous manner and could be navigated by scrolling. A dual-task methodology was applied to measuring mental workload by calculating an average reaction time to the secondary task. Results of the study demonstrated that users experienced lower mental workload when text was displayed on pages optimized to fit the computer screen and did not need to be scrolled.

**Content organization.** Organization of content in hypermedia has been found to affect learners with high and low prior knowledge differently. Well-defined structures

were shown to be more suitable for learners with low prior knowledge, because such structures help these learners clearly see the hierarchical relationships of knowledge organization that they are lacking (Potelle & Rouet, 2003). Texts that are highly coherent and thus provide more structure to the reader help low prior knowledge learners construct better situational models and stay oriented while exploring information (Salmeron, Cañas, Kintsch, & Fajardo, 2005). Caution, however, is warranted not to oversimplify material presentation. Too simple a presentation (e.g., linear, hierarchical) may significantly reduce germane cognitive load, decrease the motivation, and lead to passive learning (Antonenko & Niederhauser, 2010).

When it comes to learners with high prior knowledge, they, on the contrary, usually benefit from or are not affected by the structure and organization of information provided by hypermedia systems (Amadieu, Tricot, & Mariné, 2009). This happens primarily because these learners base their navigational decisions on their own existing mental models of the domain that help integrate new knowledge. It is, therefore, possible to assume that learners with high prior knowledge are less prone to source of cognitive load associated with navigational decisions. Such differences can be explained by the expertise reversal effect (Kalyuga, Ayres, Chandler, & Sweller, 2003), which suggests that for high prior knowledge learners, high levels of guidance provided by the system may no longer contribute to learning or even inhibit it.

**Learner control mechanisms (hyperlinks).** Another characteristic of hypermedia systems that is likely to result in fluctuations in experienced cognitive load is the number and the type of hyperlinks. DeStephano & LeFevre (2007) conducted a comprehensive overview of the literature on how decision-making with regards to which

hyperlinks to follow can impose demands on the working memory and impede comprehension and navigation. Even though the reviewed studies yielded inconsistent evidence as to the effect of the number of hyperlinks on comprehension and navigation, no evidence was found to show that more hyperlinks led to better performance. In a later study, Madrid, Van Oostendorp, & Puerta Melguizo (2009) followed up on DeStephano & LeFevre's (2007) predictions of high levels of disorientation and cognitive load caused by the greater number of hyperlinks and also found no effect of the number of hyperlinks on learning. Their additional findings suggested that the type of hyperlinks, not their number, could be a more important factor in material comprehension.

It has been found that when learners find themselves using hypermedia, they choose to follow hyperlinks because of one of the three reasons: hyperlink coherence, default position of hyperlinks on the screen, and individual interest in the hyperlink (Protopsaltis & Bouki, 2006; Salmerón, Kintsch, & Kintsch, 2010). The coherence strategy consists of selecting hyperlinks that are semantically related to the one that was just previously selected (Foltz, 1996). This strategy is the most demanding of the three because it requires the learner to engage in active reasoning and identify semantic relationships between the hyperlinks, thus, potentially increasing demands on cognitive load. At the same time, this strategy is the most rewarding as it leads to better comprehension, mostly, in low prior knowledge learners (Salmeron, Kintsch, & Cañas, 2006).

*Hyperlinks and prior knowledge.* In relation to hypertext, presence or absence of schemas or prior knowledge may influence the navigational choices made by the learner. For example, individuals with extensive prior knowledge are better able to identify their

information needs and make selections based on schema organizations in their memory (Norman, 1982). As Gall & Hannafin (1994) put it, “Increases in prior knowledge aid the individual in generating reasons to search the knowledge base, as well as establishing connections among existing and new knowledge” (p. 222). As a result, learners with high prior knowledge benefit from sequencing material for themselves and are less likely to experience cognitive overload imposed by navigational decision-making.

Individuals with low prior knowledge, on the contrary, are more likely to experience difficulties with the schema-driven selections of information as they are unable to establish information needs in advance and it is difficult for them to see the relationships between the concepts. The consequence of high cognitive demands associated with decisions as to which hyperlinks to follow may impose high cognitive demands as well.

Due to lack of situational models, learners with low prior knowledge may start following hyperlinks purely based on their interests and curiosity. Such behavior is particularly common when the learning goal calls for a broad understanding of the material and often leads to discovery and incidental learning (Heller, 1990). The contemporary theories of human information behavior support the idea of incidental learning by connecting it to encountering information through accident or discovery (Erdelez, 1999). Behavior driven browsing of the hypermedia environment, on one hand, may result in unintentional acquisition of useful knowledge; on the other hand, such behavior may be distractive to learning. Every time the learner notices interesting content and decides to proceed with learning more about it, the ongoing information acquisition

process is interrupted and there is a 50 percent chance that the learner will return to the page that triggered new informational encounter.

Interruptions in information processing have been found to impede readers' comprehension of the text because after the interruption takes place, text information needs to be reinstated in the working memory in order to allow for a successful development of a situational model (Lorch, 1993). In addition, selecting hyperlinks primarily based on learner interests may result in disorientation within the hypertext, loss of track of the reading goal or of the material activated in their working memory (DeStefano & LeFevre, 2007). When the learners find themselves in situations of not knowing how to get to a certain node or not knowing whether certain information is available within the hypertext, they are believed to be lost in hyperspace (Conklin, 1987). Loss of context of the ongoing information acquisition can, therefore, be detrimental for learning, especially for learners with low self-monitoring skills or those not familiar with the domain (Roberts & Newton, 2001) as the lack of schemas and mental models makes it easy for them to get 'carried away' in the large and complex data structures.

*Hyperlinks and cognitive overhead.* Emphasis on learner control in hypermedia environments allows freedom of navigational decisions. However, in certain instances, deciding on which hyperlinks to follow can be too demanding on the learners' working memory. Conklin (1987) referred to cognitive load imposed on the hypertext readers as a result of having to make decisions as to which hyperlinks to follow as cognitive overhead. He compared cognitive overhead to a dual-task paradigm in cognitive science, i.e., navigation planning and information retrieval occupy cognitive resources of the working memory, making them unavailable for efficient information processing.

Zumbach (2006) and Zumbach & Mohraz (2008), however, questioned the validity of the phenomenon of cognitive overhead. They elaborated that when it comes to hypermedia learning, increase in extraneous cognitive load due to non-linear information access and underlying processes of navigation planning could also lead to a deeper processing of hypertext-based information and enhance learning. They argued that learning success is higher in hypertext environments when the complexity of the material is high, because possible disorientation inherent to hypertext may be advantageous to meeting the complexity of an ill-structured domain, e.g., medicine. That is why under circumstances of complex schema acquisition, extraneous cognitive load that is typically associated with disorientation within the learning environment does not necessarily lead to decreased learning performance. Instead, a deeper and elaborated information acquisition process could be initiated. Similarly, Gerjets & Scheiter (2003) found that when learners engage in more demanding, meaningful tasks, they tend to overlook sources of extraneous cognitive load and focus more on the processes of meaning construction.

*Hyperlinks and self-regulated learning.* Emphasis on learner control also makes learning in hypertext environments a very demanding cognitive activity (Azevedo, 2005). In order for distractions not to become detrimental for learning, the learners must be able to “monitor their own comprehension of the information presented in hypermedia, select appropriate strategies for correcting any misconceptions, and develop information seeking strategies that facilitate integrating information and synthesizing information from the hypermedia” (Jonassen & Grabinger, 1990, p. 20). Evidence suggests that learners who demonstrate more advanced self-regulated learning behavior benefit more

from hypertext as they tend to develop more sophisticated mental models supported by complex network relationships between the information nodes (Azevedo, 2005).

Salmerón, Kintsch, & Kintsch (2010) evaluated the hypothesis that self-regulation may be critical to the hyperlink selection strategies, which, in turn, may lead to different comprehension gains with the help of the Winne's (1995, 2001) model of self-regulated learning. The model consists of several components of self-regulated learning, namely, (1) the importance of the learning goal, (2) the learning strategy, (3) calibration of comprehension, and (4) prior knowledge. These authors found that participants who demonstrated higher level of self-regulation (i.e., those who set high learning goal and applied appropriate learning strategies such as note taking and summarizing) were more likely to invest more cognitive effort in identifying semantic relations between information pieces. As a result, they selected links based on semantic coherence, which improved their comprehension outcomes. Further, Salmerón, Kintsch, & Kintsch found that among the readers with low prior knowledge those who calibrated their comprehension against the goals during reading also tended to select semantically related links and achieve better comprehension.

When it comes to hypermedia learning, presence of self-regulated learning and self-monitoring skills is a definite advantage. These skills help coordinate cognitive processes and facilitate constructivist learning as a result of learner-driven navigational decisions. At the same time, concerns are raised that active utilization of these skills may occupy the limited capacity of the working memory and interfere with meaningful information processing (Schwartz, Andersen, Hong, Howard, & McGee, 2004). As a result, it is suggested that instructional designers use recognizable structures of

hypermedia environments in order to reduce the metacognitive demands on the working memory that otherwise are spent on deciphering unfamiliar information structures (Schwartz et al., 2004).

### **The Cognitive Theory of Multimedia Learning in Hypermedia Learning**

It has been suggested that established theories of multimedia learning may be useful in explaining how the design of hypermedia environments can affect cognitive load. Dillon & Jobst (2005) proposed that Richard Mayer's Cognitive Theory of Multimedia Learning (2005) is “the only dedicated theory of multimedia learning that explicitly aims at guiding our analysis and understanding of learning in hypermedia environments” (p. 570). The Cognitive Theory of Multimedia Learning (CTML) similar to Cognitive Load Theory emphasizes the sensitivity of human cognitive architecture to cognitive overload in electronic environments due to shortcomings of instructional design. Both theories prescribe how different ways of content presentation (i.e., auditory and visual) should be combined in order to foster meaningful learning and keep extraneous cognitive load as low as possible. That is why the CTML not only explains how people learn from multimedia instruction but also points at the sources of cognitive overload in hypermedia environments.

The fundamental principle of multimedia learning states that people learn from words and pictures better than from words alone (Mayer, 2005). This is explained by the fact that presentation of information in dual channels enables the construction of two separate mental models of knowledge representation – pictorial and verbal (cf., Baddeley, 1992), which increases the likelihood of information to be effectively integrated into long-term memory without being lost. Multimedia instruction that includes animation can

improve learning even more. By helping students visualize the process, animations have been found most effective when presenting information that students may have difficulty imagining due to lack of familiarity with the content or the concepts (Betrancourt, 2005). In addition, animations are likely to be more effective when the learners can interact with them and view them at their preferred pace (Mayer & Chandler, 2001).

Despite the benefits of dual information processing, it may happen that pictures or animation are presented separately from the textual description accompanying them. In such situations learners' attention is split because they need to attend to information that is spatially or temporally distant. Sweller et al. (1998) referred to this phenomenon as a split-attention effect. Forced integration of information from two or more verbal and pictorial information sources increases demands on the working memory, thus, causing cognitive overload and inhibiting learning. As opposed to the established opinion that split-attention effect is caused primarily by excess of extraneous cognitive load, Cierniak, Scheiter & Gerjets (2009) demonstrated that both extraneous and germane cognitive load may contribute to cognitive overload during split-attention effect.

Multimedia learning is more effective when the learner has the ability to interact with material presentation and control its pace (Mayer, Dow, & Mayer, 2003). In situations when the learner is unable to meet the required processing demands, cognitive overload is likely to occur. A potential solution to this problem is breaking presentation into shorter, more manageable segments and allowing some time between these segments (Mayer & Moreno, 2003). Although it is a common assumption that complexity of material cannot be altered by instructional design, segmenting has been proposed as a technique for simplifying material presentation by allowing learners to process

information one piece at a time (Mayer & Moreno, 2003). Enabling the learner to control the pace of material presentation helps with managing the available cognitive capacity and preventing cognitive overload.

Learning in multimedia is also most effective when it includes only relevant content (Mayer, 2003). According to CTML, adding potentially interesting but extraneous information that does not directly support learning may leave the learner with the limited cognitive capacity for meaningful information processing (Mayer, 2005). This type of extraneous material in an instructional presentation is referred to as extraneous cognitive load. Mayer & Moreno (2003) proposed to address problems like this one by a load-reducing technique called *weeding*, which consists of supplying the learner only with essential information in a short and concise manner.

Finally, multimedia learning is more effective when the structures for organizing information are activated prior to exposure to content (Pollock, Chandler, & Sweller, 2002). Mayer & Moreno (2003) proposed *pre-training* and *signaling* as strategies for knowledge activation. Pre-training effect provides the learners with an opportunity to get familiar with system components before they get to interact with the system. The purpose of the signaling effect is to present the learners with the signals concerning how to process the material in order to activate their prior knowledge, which should prove helpful in selecting and organizing relevant information.

Along the same vein, in an attempt to improve learning in hypermedia, Hoffman (1997) proposed to rely on the Elaboration Theory (Reigeluth & Stein, 1983) for content organization and presentation. According to this theory, instructional material needs to be presented in a certain sequence, beginning with the special kind of overview that presents

a few general but fundamental ideas about the remainder of the instruction. Hoffman (1997), therefore, proposed that in accordance with Elaboration Theory, hypermedia learning environments should include overviews available to learners at the very beginning, which should help them make decisions with regards to selection of information by providing access to content organization.

Despite the attempts to incorporate the principles of the CTML to learning in hypermedia, it was noted that applying this theory to evaluating learning with hypermedia may be too simple as following its principles did not result in any noticeable improvements in learning (Gerjets et al., 2009). Nevertheless, I choose to believe that the principles of multimedia learning can be informative to hypermedia learning for a few reasons. First, hypermedia is a special case of multimedia (Kommers et al., 1996), which means that hypermedia design may include a number of multimedia elements. As a result, certain design principles applicable to multimedia environments are applicable to hypermedia as well. Second, CTML can be informative in understanding how design of hypermedia environments that contain multimedia elements can affect cognitive architecture of learners.

### **Measurements of Cognitive Load**

Cognitive Load Theory (CLT) is one of the fundamental theories to describe the cognitive processes occurring during learning in electronic environments (Mayer, 2005). CLT explains the effectiveness of these environments for learning from the theoretical perspective; therefore, it is extensively used to inform their instructional design. The key concept of CLT is cognitive load, which describes internal processes of information processing that are not directly visible. In order to use this concept to inform design of

the learning material, data about valid and reliable assessment of cognitive load are needed (Brünken et al., 2010).

Paas & Van Merriënboer (1994) originally proposed that cognitive load is a multidimensional construct that consists of three measurable dimensions of mental load, performance, and mental effort. They defined mental load as the aspect of cognitive load that provides an indication of the expected cognitive capacity demands imposed by the task. By providing an expected estimate of task demands, mental load can be used as an a priori estimate of cognitive load. Performance was defined as the aspect of cognitive load that is demonstrative of learner achievements, e.g., the number of errors, the number of correct answers, time on task, etc. Performance can be assessed while learners are working on the task as well as after task completion. Finally, mental effort was defined as the aspect of cognitive load that refers to the cognitive capacity actually allocated to accommodate the demands imposed by the task. Because of the nature of mental effort, this aspect of cognitive load can be measured only while the learner is working on a task. Paas & van Merriënboer (1994) stated that mental effort provides the most reliable estimate of cognitive load because it shows the actual intensity of the cognitive effort that the learner expends while working on a task. They believed that mental effort might yield important information that cannot be reflected in mental load or performance.

Following Paas & Van Merriënboer's (1994) model, it is possible to separately measure each of its three dimensions of cognitive load. Estimates of mental load can be established through analytical methods that include mathematical models, task analysis, and subjective techniques such as expert opinion (Xie & Salvendy, 2000). Estimates of mental effort and performance can be established through empirical methods and may

include 1) rating scales based on subjective opinion of learners; 2) performance data (e.g., test scores); 3) physiological data (e.g., increased heart rate); and 4) psychophysiological data that combines physiological measures with behavioral characteristics (e.g., eye-tracking) (Paas & van Merriënboer, 1994).

Brünken et al. (2010) suggested classifying methods for measuring cognitive load based on their objectivity and causal relationship. Self-reported measures of cognitive load were suggested to be classified as subjective as opposed to objective measures that could include physiological, psychophysiological, behavioral, and learning outcomes measures. On the other hand, physiological, behavioral, and learning outcome measures were classified as indirect because they are indirectly linked to cognitive load and could also be the result of factors such as attention or motivation (Brünken, Steinbacher, Plass, & Leutner, 2002). Direct measures include direct measures of cognitive processes during task execution and are best used to visualize the detailed trend and pattern of cognitive load. Among them are positron-emission tomography and functional magnetic resonance imaging that allow measuring brain activation during task execution (Smith & Jonides, 1997); dual-task-paradigm that is well known in experimental psychology (Verwey & Veltman, 1996); and task-evoked pupillary response (TEPR) (Beatty & Lucero-Wagoner, 2000). Due to their directness and objectivity, physiological and psychophysiological measures of cognitive load are becoming very promising techniques that carry the potential of making important contribution to the general framework of research on cognitive load, and, in particular, hypermedia learning (Paas et al., 2008).

### **Eye Tracking as a Psychophysiological Measure of Cognitive Load**

Eye tracking technology, as the term suggests, is used for tracking and recording individual's eye movement. It is an informative methodology that offers a look inside people's visual and cognitive processes. By providing objective data of these processes, eye tracking has been used for diagnostic purposes as well as for studying human-computer interaction (Duchowski, 2002). As a result, eye movement has been informative for a large scientific audience including clinicians, vision researchers, transportation experts, usability evaluators, educators, linguists, and those who are interested in measures of cognitive processing (see overview in Duchowski, 2003).

**Factors Affecting Eye Movement in Hypermedia Learning.** Learning in hypermedia involves interacting with and extracting information from text and images. In the case of electronic information resources that I view as examples of hypermedia learning environments, the majority of information is available via text, which makes reading the primary method of information acquisition and processing.

*Physiological processes of eye movement.* Reading is a higher order cognitive activity, which is possible due to decoding of letters and words as well as establishing connections between ideas or propositions that occur in the text (Wittrock, 1990). On the other hand, reading is possible due to mechanics of eye movement (Rayner, 1998). During information processing, be it reading, visual search, or scene perception, the two basic components of eye movements are saccades and fixations (Rayner, 2009). During fixations, the eye remains fairly still and new information is acquired from the visual array. Saccades, on the other hand, are fast ballistic movements of the eye that occur between fixations. Under most normal circumstances, new information cannot be

obtained during saccades because saccades are characterized by a rapid eye movement, during which only a blur could be perceived (Uttal & Smith, 1968).

*Lexical, syntactic, and semantic characteristics affecting eye movement.* Reading also becomes possible due to psycholinguistic processes that underlie written language comprehension (Rayner, 1998). As a result, lexical, syntactic, semantic, and typographical characteristics of the text can impact eye movement and the quality of cognitive processes required for information processing. For example, as word length increases, the probability of fixating (and re-fixating) on that word increases as well (Rayner & McConkie, 1976). Words of two-three letters are fixated around 25 percent of the time, whereas words of eight letters and longer are almost always fixated (Starr & Rayner, 2001). A number of lexical, syntactic, and discourse variables have also been found to affect time of fixation duration. For instance, readers spend more time fixating on lower frequency words (e.g., unfamiliar or complex words), which are less likely to be encountered during reading, and less time fixating on higher frequency words (Rayner et al., 1996). In addition, when words are easy to predict from previous context, they are fixated on for shorter duration (Inhoff & Rayner, 1986).

The spatial and temporal features associated with fixations and saccades can be measured and are both of considerable interest. These measurements include fixation duration, fixation frequency, saccade landing position, regressive saccades (e.g., backtracking eye movement), gaze duration, total viewing time, scanpaths (e.g., a complete saccade-fixate-saccade sequence), blink rate, and pupil size (see overview in Poole & Ball, 2005). For instance, as text becomes conceptually more difficult, fixation duration increases, saccade length decreases, and the frequency of regressions increases

(Jacobson & Dodwell, 1979). Fixation frequency in a certain area of interest indicates the degree of importance and fixation duration indicates the complexity and difficulty of visual display (Fitts, Jones, & Milton, 2005). When the readers initially misanalyse syntactically ambiguous sentences, they tend to spend more time on the disambiguating word and make regressions in order to re-read the sentence (Paterson, Liversedge, & Underwood, 1999; Rayner, Carlson, & Frazier, 1983). When the readers are experiencing high workload, their blink rate goes down; when they are fatigued – the blink rate goes up (Bruneau et al., 2002).

Fixations also serve as a measure of attention and attentional shifts, be it attention triggered by the external (e.g., environmental factors) or internal stimuli (e.g., cognitive processes imposed by the demands of the task) (Underwood, Chapman, Brocklehurst, Underwood, & Crundall, 2003). Further, Togami (1984) experimentally showed that the number of fixations and fixation duration corresponds to the accuracy of visual inspection. This author explained that longer fixation duration leads to improved performance, possibly due to increased levels of attention paid to details available in the content.

Another characteristic of eye movement unrelated to fixations and saccades – changes in pupillary response – could also be informative in reading research. Cognitive pupillary response has been recognized as a valid physiological measure of mental effort or processing load (Beatty & Lucero-Wagoner, 2000; Beatty, 1982; Iqbal, Zheng, & Bailey, 2004) and as an indicator of information processing activity during reading (Carver, 1971). A number of authors reported that mean pupil size, maximum percentage of pupil dilation, and pupillary lability reflect difficulty of the material specifically on

complex reading tasks (see overview in Metalis et al., 1980). Iqbal et al. (2004) showed that pupillary response correlated with the workload of interactive tasks and discovered that changes in workload align well with fluctuations of pupil dilations. Poock (1973) suggested that pupillary diameter could be used as an indicator of mental overload because in his study students demonstrated severe constrictions of the pupil when they were required to process information at a rate above their maximum capacity. Beatty & Lucero-Wagoner (2000) summarized a few studies and suggested that task evoked pupillary response amplitude can also be an appropriate quantitative measure that adequately reflects the state of central attentional processes.

The fact that (1) the pupil tends to dilate relative to baseline during the performance on cognitive tasks, (2) the degree of maximum dilation on a given task is a reliable indicator of task difficulty for the subject, and (3) the moment-to-moment relative size of the pupil reflects the total momentary cognitive processing load suggest that the pupil may reflect cognitive processing not only in simple problem tasks but also in complex continuous tasks such as reading (Metalis et al., 1980).

It is also worth noting that despite the influence of spatial and temporal features of the text and reading environment, individual differences affect eye movement. For example, less skilled readers (e.g., beginner, poor, dyslexic) take longer to read, which is reflected in typically have longer and more frequent fixations and shorter saccades (Starr & Rayner, 2001). Fast readers, on the other hand, demonstrate shorter fixations, longer saccades, and fewer regressions (Underwood, Hubbard, & Wilkinson, 1990). Bilingual readers make shorter fixations, longer saccades, and fewer regressions in their dominant language (Altarriba, Kroll, Sholl, & Rayner, 1996). As learners' expertise and knowledge

base increase, they tend to fixate (faster) on more task-relevant information (Haider & Frensch, 1999).

*Interface design characteristics affecting eye movement.* In addition to lexical, syntactic, and individual characteristics, readers' eye movement representative of cognitive processes can be affected by interface design, e.g., color schemes, font type and size, content organization, images and animation. A few studies demonstrating this effect are described below.

Bhattacharyya, Chowdhury, Chatterjee, Pal, & Majumdar (2014) investigated the effect of character/background color combination on eye movement, cognitive load, and performance data during onscreen searching tasks. They found that light character color on a dark background interfered with retinal image formation and had a significant effect on pupil diameter, mean fixation duration, and fixation frequency. The authors, however, were not sure whether these changes were due to the increase in cognitive load associated with the color combinations or by the dark color background itself. As a result, light backgrounds were named as preferable as long as they were not combined with light-colored objects. Among other findings, the authors reported that blue was advantageous as background or as character and associated with lower cognitive load, lesser searching time, minimum error and optimum eye movement activities.

Hyönä & Lorch (2004) examined the effects of topic headings on processing of multiple-topic expository texts with the help of eye fixations. They found that the presence of topic heading attracted more fixations and, thus, facilitated processing of topic sentences (i.e., the sentences that introduce a new topic in the text) and increased the number of topics mentioned in the text summaries written after reading the text. This

finding suggested that topic headings could function as signals, especially when readers are actively looking for a relevant topic. As soon as the new topic was identified, a new memory slot opened for it and a mental frame was starting to build around it with subsequent information integration.

Leuthold, Schmutz, Bargas-Avila, Tuch, & Opwis (2011) applied eye tracking technology to investigating the influence of menu design on user performance, navigation strategy, and personal preferences. Two types of menus were compared: vertical, where users only had to process the number of visible menu items, and dynamic, where users needed to click on the group title to display the navigation items it contained. Findings from this study demonstrated that when users were presented with vertical menus, they needed to fixate less on the menu before making a decision as to what menu item to select, made those decisions significantly faster, and rated their interaction with vertical menus as easy, less frustrating, and more helpful.

Holsanova, Holmberg, & Holmqvist (2008) analyzed user eye movements as they read newspaper with different design of information graphics and related text. They found that a graphic design where related text was presented near the graphic rather than far from it made it easier to integrate information from two information sources. This was confirmed by a higher number of integrative saccades that represented transitions between semantically related pieces of verbal and pictorial information and indicated the process of reader making connections between text and illustrations.

Goldberg (2012) studied how perceived web page complexity related to eye movement metrics. Participants were presented with 25 different variations of the stimuli and were asked to rate their complexity on a scale from 1 (extremely overwhelming) to 7

(extremely clear). The results showed that pages that were rated as less visually complex (i.e., pages that required less time viewing, contained fewer page elements, and were less dense in content and color) were associated with longer fixation durations. Goldberg (2012) explained that longer fixations were the result of higher attention because users could focus the attention easier on the pages with less visual clutter.

Eye movement studied through the combination of the described metrics can be an excellent behavioral measure of the cognitive processes underlying reading (Liversedge & Findlay, 2000; Rayner, 1998). Eye movement as a cognitive measure, however, should be approached with caution because typographical characteristics of text such as font variations, line length, letter spacing (Kolers, Duchnicky, & Ferguson, 1981; Morrison & Inhoff, 1981) as well as the factors of the surrounding environment (e.g., ambient light levels, display color) (Goldberg & Wichansky, 2002) may also influence eye movement.

Despite the presence of research evidence on applying eye tracking to studying the processing of multimedia material, these studies are still in short demand (Hyönä, 2010; van Gog & Scheiter, 2010). The main reasons for scarcity of research implementing eye tracking methodology is very rich data that requires challenging and time-consuming analysis. Nevertheless, because eye tracking is informative in providing accurate estimates of cognitive processes, there is a need for more studies in the direction that could help interpret hypermedia effects on and during learning.

## **CHAPTER 3 – METHODS**

### **Research Objective and Questions**

The purpose of the study was to explore the appropriateness of three electronic information resources for learning by measuring learners' mental effort invested in interactions with these resources. The research goal was achieved by seeking answers to the following research questions:

1. What are the differences in learners' mental effort invested in interactions with electronic information resources?
2. How do the differences in learners' mental effort invested in interactions with electronic information resources affect learning?
3. How do certain interface design characteristics of electronic information resources contribute to or interfere with information acquisition and, consequently, learning?

### **Study Setting and Participants**

The study took place during the summer semester of 2013-2014 academic year. Participants of the study were 18 first-year medical students representing Class 2017 of the University of Missouri School of Medicine (MU SOM) who started their program in August, 2013. By the time of the experiment, these students had successfully completed their first year of pre-clerkship curriculum but had not yet started their second year.

A recruitment announcement was sent via a mass email targeting the medical students of Class 2017. To encourage participation, the recruitment email was sent three times throughout the month of July, 2014. Students were invited to participate in the study voluntarily and offered monetary compensation for their time in the amount of \$50

per session. Due to the nature of the study, students who suffered vision loss; had ocular prosthesis; wore prescription glasses of 4+ diopters; bifocal, trifocals, progressive, and transition lenses were not eligible for participation.

## **Materials and Instruments**

The following materials were prepared for the experiment.

### **Questionnaires**

1. A demographic questionnaire to obtain demographic characteristics of participants (Appendix A).
2. A questionnaire on students' preferences for electronic information resources for learning (Appendix B).

### **Learning Material**

A second-year medical student of Class 2016 of the MU SOM transitioning into his third year of medical school provided assistance with identifying learning topics. This student's research interests and research experience were in medical education, and he was one of the highest achieving students in his class. Having gone through the same curriculum, this student was aware of the learning material covered and not covered during the first year of school and could approximate the scope of medical knowledge required from students upon completion of the first year. He proposed to include the following three diseases as learning topics: Idiopathic Pulmonary Fibrosis (IPF), Primary Sclerosing Cholangitis (PSC), and Neurosarcoidosis (NS). These diseases were later vetted for the experiment by the MU SOM curriculum directors based on the following criteria:

1. The diseases were not covered in the first-year curriculum (including

- lectures, PBL tutorials, or exams).
2. The diseases had extensive coverage in all three electronic information resources.
  3. Despite the fact that the diseases described different body systems, they were similar in their complexity.
  4. The diseases were rare, which ensured the complexity of the underlying processes of disease mechanisms. At the same time, the mechanisms of these diseases should have been intellectually accessible to study participants as they required the pool of medical knowledge gained during the first year of program.

Including novel and difficult learning material in the experiment was critical. In accordance with the cognitive load theory, because intrinsic and extraneous cognitive loads are highly additive (van Merriënboer & Sweller, 2010), extraneous cognitive load becomes an issue when intrinsic cognitive load of the learning material is high. That is why, to understand whether or not the differences in information presentation and organization in three electronic information resources interfered with proper information acquisition, I opted for the learning material to be intentionally complex.

### **Screening Test**

The purpose of the screening test was to obtain students' self-assessment of their familiarity with and perceived medical knowledge of the diseases (Appendix C). Students' familiarity with the diseases included their informal knowledge of these diseases (e.g., I heard about it, I read about it, or I had an acquaintance diagnosed with this medical condition) as opposed to purposefully obtained medical knowledge (e.g., I

independently and purposefully studied the diseases – the mechanisms, pathophysiology, or management).

## Knowledge Test

Three knowledge tests were composed for the experiment (Appendix D). Each knowledge test intended to assess students' understanding of each disease through ten multiple-choice questions. The test questions were generated from primary medical literature, namely, peer-reviewed review articles found on each disease (Appendix D). Structurally, questions were formatted in accordance with the manual on constructing written test questions for the basic and clinical sciences (Case & Swanson, 2002). That is, each question was a two-step question, where step 1 included a narrative about the contextual situation to which the question applied and step 2 was the actual question (Figure 2).

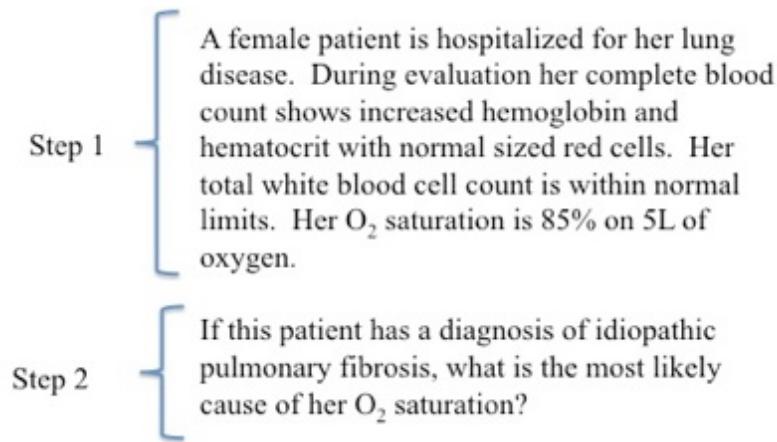


Figure 2. Two-step question organization

Students were offered seven multiple-choice options, i.e., one correct, four distractors, and two more answer options of “*I don't know*” and “*I can't remember*

*reading about it*" to prevent participants from random guessing. The knowledge tests were offered to students in the electronic format created in Qualtrics© online survey software. During the administration of the tests, the order of test questions on each disease was automatically randomized for each participant.

**Content validity of the learning material and the knowledge test.** The first step in validating the learning material and the knowledge test included meeting with two representatives of the Pre-clerkship Curriculum Steering Committee of the MU SOM in order to vet the appropriateness of the diseases and test questions from the learning standpoint. The first reviewer held a PhD in Pharmacology and served as a Faculty Director of Evaluation at MU SOM. The second reviewer held a PhD in Microbiology and served as a Faculty Director for Basic Sciences Education at MU SOM. Both reviewers had extensive experience writing, reviewing, and validating learning materials for examination and testing practices in pre-clerkship curriculum.

During a collaborative process, the reviewers assessed the chosen learning material for relevance and confirmed that it met the above-mentioned criteria (see pp. 45-46 above). Further, each question and five multiple-choice answer options of the knowledge tests were thoroughly reviewed for clarity of meaning and structure as well as relevance. The reviewers recommended revising a few questions due to their ambiguity, which was successfully accomplished.

The second step included validating the knowledge test with physicians, residents, and senior medical students in order to confirm the correctness of the identified answers and their availability in the three electronic information resources. Three reviewers from the University of Missouri – one physician, one resident, and one third-year medical

student – were each assigned an information resource (e.g., either *UpToDate*, *Access Medicine*, or *Wikipedia*) and instructed to seek answers they deemed correct to the knowledge tests on three diseases. These reviewers were specifically asked not to rely on their prior knowledge in answering the test questions and were not given any time limitations.

After the correct answers were identified, they were compared with the correct answers identified by the 2<sup>nd</sup> year medical student who helped design the experimental materials. The overall initial reviewer agreement per disease and per information resource was satisfactory (Table 2). Disagreements were resolved through discussion resulting in 100 percent final agreement. Changes to the wording of some questions suggested by the reviewers were implemented.

Table 2

*Initial Reviewer Agreement on the Knowledge Test Answers*

	Idiopathic Pulmonary Fibrosis	Primary Sclerosing Cholangitis	Neurosarcoidosis	Overall agreement per information resource
Reviewer 1 ( <i>UpToDate</i> )	80%	80%	60%	73%
Reviewer 2 ( <i>Access Medicine</i> )	100%	70%	80%	83%
Reviewer 3 ( <i>Wikipedia</i> )	100%	40%	80%	73%
Overall agreement per disease	93%	63%	73%	

**Face validity of the learning material and the knowledge test.** Face validity of the learning material and knowledge tests was established through the pilot test with one first-year medical student who represented the population of interest. The pilot test lasted approximately two hours and the participant received \$50 for participation. The goal of the pilot test was to identify any opportunities for improvement to the instruments and to confirm the assigned time duration for the experiment. During the pilot test the student was asked to go through all steps of the experimental procedure and comment if anything was unclear. The major points in the feedback from the participant are summarized in Table 3.

Table 3

*Pilot Test Feedback*

<b>Feedback</b>	<b>Actions taken</b>
It was confirmed that the three diseases chosen for the experiment were not included in the first-year curriculum.	n/a
Instructions about the task needed to be more clearly defined. The first version said: "Please read and learn about the disease. You will be tested on your knowledge of pathophysiology, physiology, treatment, and diagnosis of this disease".	Instructions were revised to "Please read and learn about the disease. <i>Read the text thoroughly and pay attention to details.</i> " You will be tested on your knowledge of pathophysiology, physiology, treatment, and diagnosis of this disease".

It was suggested to have a fair warning about the nature and the number of the questions in the knowledge test.

It was suggested to provide participants with specific links in information resources to learn about diseases instead of allowing students to decide on the source on their own.

It was found that one of the diseases (i.e., idiopathic pulmonary fibrosis) was easier to learn about than the other two.

Instructions included the mention that each comprehension test consisted of ten multiple-choice questions and the question format would be similar to that of end-of-block and end-of-year exam questions, e.g., clinical, two-step.

This suggestion was rejected because I did not want to limit students' personal choices of information resources.

The participants could have had such a perception due to the nature of the information resource assigned to read and learn about the disease. To take this into account, an additional question "Which disease was the easiest and the hardest to read and learn about?" was added to the interview protocol.

---

## **Apparatus**

Three pieces of technology were instrumental for this study: Tobii eye-tracker, Morae recording software, and a computer workstation. Below is the description of how exactly the technology was used during the experiment.

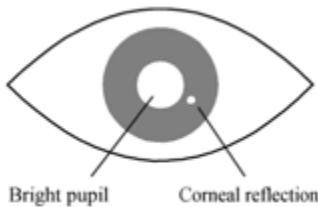
### **Tobii Eye Tracker**

Technology that allows to track eye movement is often referred to as eye tracker. In this study I used Tobii X2-60 Eye Tracker by Tobii corporation (<http://www.tobii.com>). By being a modular eye tracker with a single USB cable that can be mounted on a desktop, laptop, or a tablet, this particular model is the smallest eye-tracking research system on the market (Figure 3).



*Figure 3. Tobii X2-60 eye tracker*

Like other commercial eye-tracking systems available today, Tobii measures point-of-regard by the “corneal-reflection/pupil-center” method (Goldberg & Wichansky, 2002). It uses the center of the pupil and infrared light to create corneal reflections (Figure 4). The light enters the retina and a large portion of it is reflected back, making the pupil appear to be bright or dark, depending on the location of the illumination source. Tobii X2-60 has dual camera system that automatically selects bright or dark pupil, hence, allowing for variability in experimental conditions.



*Figure 4.* Corneal reflection and bright pupil

Tobii X2-60 captures data at a temporal resolution of 60 Hz, which means that 60 samples of an eye are captured per second. Tobii eye-tracker allows tracking of both eyes, which increases the accuracy and reliability of data and ensures very low data loss. This eye-tracker also demonstrates tolerance to a large freedom of head movement, which means that participants can move freely and naturally during the experiment. Like most video-based eye-trackers, Tobii X2-60 needs to be fine-tuned to the particularities of each person's eye movements through the process called calibration. Calibration is performed at the beginning of data collection by displaying a dot in different parts of the screen and analyzing eye response.

After the eye movements are recorded, the data becomes available in the numeric form through Tobii Studio, the eye tracking analysis and visualization software. Available data also include (1) animated representations of a point on the interface, (2) static representation of the saccade path, (3) heat maps, and (4) blind zone maps.

### **Morae**

*Morae Recorder* (version 3.3.1) distributed by TechSmith is a recording software that allows capturing real-time user interactions with the content delivered through the computer interface (including voice recordings). *Morae Recorder* was used to capture

participants' interactions with electronic information resources. Another component – *Morae Observer* was used to enable remote observations of user interactions with information resources.

### **Computer workstation**

Both Tobii eye-tracker and *Morae* were installed on a Dell Latitude E6530 computer workstation with a 15.6" display. This workstation was located in a small isolated room with constant artificial lighting projected from the ceiling.

### **Procedure**

Data collection took place during the summer semester of 2013-2014 academic year after obtaining the approval #1211506 of the MU Institutional Review Board. Data collection took place on the premises of the MU campus. The experiment was conducted in a controlled laboratory setting, with one participant being in a room at a time. All data collection sessions were conducted in the same room.

Step 1. After each participant was informed about the details and procedures of the study, s/he was asked to sign a consent form and complete a demographic survey and a survey on preferences for information resources.

Step 2. Each participant was asked to find a comfortable sitting position in front of the computer monitor at a distance of approximately 60-65 cm (24-24 inches) between his/her eyes and the eye-tracker. After a comfortable position was found, calibration was performed. Participants' eye movement was recorded from this point forward.

Step 3. Each participant was asked to take a screening test to establish his/her familiarity and medical knowledge of a disease.

Step 4. The main part of the experiment consisted of each participant reading about three diseases from three information resources. The first page the participants saw on the computer monitor was the page with the search results for each of the medical conditions in the assigned information resource. From that page, the participants could follow any hyperlinks and choose any articles, books, chapters or pages to read about pathology, pathophysiology, symptoms, and treatment of each disease as long as they stayed within the same information resource. Reading outside the information resource was not allowed. Each participant was given about 20-25 minutes per disease<sup>3</sup> and during reading allowed to take electronic notes in a Word document. The order of information resources and diseases was assigned to each participant in accordance with the 3x3 repeated Latin square design described below.

A Latin square design (LSD) is a special type of comparative design used where there is a single factor of primary interest, typically called the treatment factor, and two nuisance factors, i.e., factors adding to variability between the treatments (Natrella, 2010). Because LSD is a comparative design, there are usually  $n$  variations of the main treatments that are being compared with one another. That is why, an LSD is represented by an  $n \times n$  table (e.g., 3 x 3, 4 x 4, or 5 x 5), where each cell of the table signifies treatment variations that occur exactly once in each row and each column (Montgomery, 2001). In this study, the treatment factor was associated with the type of electronic information resources (the cells); two nuisance factors were the participants (the rows) and the diseases (the columns) (Table 4).

---

<sup>3</sup> de Jong (2010) talked about a substantial portion of research on cognitive load theory that relies on a short study time (e.g., 6 min). He called such study time too short for a realistic learning task and emphasized the importance of realistic study times. In this study, based on the results of the pilot test, 20-25 minute study time per disease was considered realistic.

Table 4

*Example of 3 x 3 Latin Square*

<b>Participants</b>	<b>Diseases</b>		
	IPF	PSC	NS
P1	<i>Wikipedia</i>	<i>Access Medicine</i>	<i>UpToDate</i>
P2	<i>Access Medicine</i>	<i>UpToDate</i>	<i>Wikipedia</i>
P3	<i>UpToDate</i>	<i>Wikipedia</i>	<i>Access Medicine</i>

LSD was chosen for this study because it (1) handles cases with several nuisance factors that act as sources of variability appearing between successive observations of treatments, and (2) allows experiments with a relatively small number of runs (e.g., in 3 x 3 design there are nine runs) (Natrella, 2010). In case where the number of treatment variations is small (e.g., n=3) but considerable variability among them is expected, two or more Latin squares may be used to increase degrees of freedom for experimental error (“Replicated Latin Squares, STAT 503 Design of Experiments,” 2015). Altogether, there are 12 possible replications of 3 x 3 Latin squares resulting from rows and columns permutations (Figure 5).

X Y Z	X Y Z	X Z Y	X Z Y
Z X Y	Y Z X	Y X Z	Z Y X
Y Z X	Z X Y	Z Y X	Y X Z
Y Z X	Y Z X	Y X Z	Y X Z
X Y Z	Z X Y	Z Y X	X Z Y

Z	X	Y		X	Y	Z		X	Z	Y		Z	Y	X
Z	X	Y		Z	X	Y		Z	Y	X		Z	Y	X
Y	Z	X		X	Y	Z		X	Z	Y		Y	X	Z
X	Y	Z		Y	Z	X		Y	X	Z		X	Z	Y

Figure 5. Twelve possible replications of 3 x 3 Latin squares

Latin squares included in the experiment were finalized in the following ways:

- 1) To avoid the order bias effect when assigning the sequence of the diseases, six possible permutations of three diseases were identified (Figure 6):

Permutations, n	Diseases		
1	IPF	NS	PSC
2	IPF	PSC	NS
3	PSC	NS	IPF
4	PSC	IPF	NS
5	NS	IPF	PSC
6	NS	PSC	IPF

Figure 6. Permutations of the three diseases

- 2) Twelve possible 3 x 3 Latin squares were generated in random order in IBM SPSS Statistics 19 software.
- 3) Because the final sample size of participants consisted of 18 students, I ended up using only six out of twelve possible Latin squares chosen at random with the help of random number generator (<https://www.randomizer.org>) (Table 5).

Table 5

*Six Latin Squares Used in the Experiment*

Latin square #1

<b>Participants</b>	<b>Diseases</b>		
	IPF	NS	PSC
P2	<i>Wikipedia</i>	<i>UpToDate</i>	<i>Access Medicine</i>
P7	<i>Access Medicine</i>	<i>Wikipedia</i>	<i>UpToDate</i>
P12	<i>UpToDate</i>	<i>Access Medicine</i>	<i>Wikipedia</i>

Latin square #2

<b>Participants</b>	<b>Diseases</b>		
	IPF	PSC	NS
P17	<i>Wikipedia</i>	<i>Access Medicine</i>	<i>UpToDate</i>
P18	<i>Access Medicine</i>	<i>UpToDate</i>	<i>Wikipedia</i>
P19	<i>UpToDate</i>	<i>Wikipedia</i>	<i>Access Medicine</i>

Latin square #3

<b>Participants</b>	<b>Diseases</b>		
	PSC	NS	IPF
P3	<i>Access Medicine</i>	<i>UpToDate</i>	<i>Wikipedia</i>
P8	<i>Wikipedia</i>	<i>Access Medicine</i>	<i>UpToDate</i>
P14	<i>UpToDate</i>	<i>Wikipedia</i>	<i>Access Medicine</i>

Latin square #4

<b>Participants</b>	<b>Diseases</b>		
	PSC	IPF	NS
P6	<i>UpToDate</i>	<i>Access Medicine</i>	<i>Wikipedia</i>
P11	<i>Wikipedia</i>	<i>UpToDate</i>	<i>Access Medicine</i>
P15	<i>Access Medicine</i>	<i>Wikipedia</i>	<i>UpToDate</i>

Latin square #5

<b>Participants</b>	<b>Diseases</b>		
	NS	IPF	PSC
P4	<i>UpToDate</i>	<i>Wikipedia</i>	<i>Access Medicine</i>
P10	<i>Wikipedia</i>	<i>Access Medicine</i>	<i>UpToDate</i>
P16	<i>Access Medicine</i>	<i>UpToDate</i>	<i>Wikipedia</i>

Latin square #6

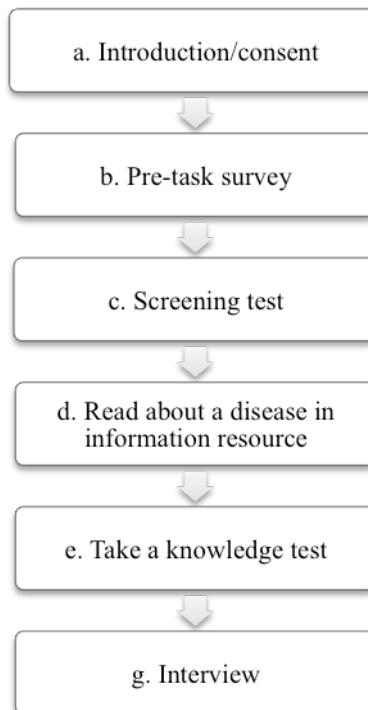
<b>Participants</b>	<b>Diseases</b>		
	NS	PSC	IPF
P5	<i>Access Medicine</i>	<i>Wikipedia</i>	<i>UpToDate</i>
P9	<i>Wikipedia</i>	<i>UpToDate</i>	<i>Access Medicine</i>
P13	<i>UpToDate</i>	<i>Access Medicine</i>	<i>Wikipedia</i>

Step 5. After the student completed reading, s/he was taken to the page with the knowledge test. While taking the test, the participants were not allowed to consult the information resource or their notes. They were also not allowed to move back and forth

between the test questions. After completing the knowledge test, participants could take a few-minute break before beginning to read about the next disease in the next information resource. To ensure the accuracy of eye movement recording, calibration was performed before each task.

Step 6. After the participant completed working on the three tasks, s/he participated in individual interview (Appendix E). The purpose of each interview was to gain understanding of students' individual perceptions of interacting and learning with electronic information resources.

An overview of the experimental procedure is shown in Figure 7. Steps d. and e. were repeated three times (i.e., students read about three diseases in three information resources).



*Figure 7. Overview of the experimental procedure*

## **Data Analysis**

### **Unit of Analysis**

A unit of analysis was each student's interactions with each information resource.

In total, 54 student interactions with information resources were analyzed (18 subjects x 3 resources). Each information resource was used six times to read about each of the three diseases. During each interaction with information resources, of interest were overall information behavior and, in particular, students' eye movement characteristics.

### **Metrics of Interest**

The following eye metrics were of interest (as defined in *User Manual — Tobii Studio* (2012) and elsewhere):

- 1) Total fixation duration (measured in seconds) – the sum of duration of all fixations within a predefined area of interest (AOI).
- 2) Fixation count (measured in counts) – the number of times the participant fixated on a particular AOI.
- 3) Total visit duration (measured in seconds) – the total amount of time that the participant spent within a predefined AOI while it was active.
- 4) Task evoked pupillary response (TEPR):
  - a. mean pupil dilation (MPD) (i.e., the average pupil diameter over a given interval of time) (Beatty & Lucero-Wagoner, 2000);
  - b. lability (i.e., the difference between the maximum and the minimum values of pupil size during the task) (Matalis et al., 1980);
  - c. percentage change in pupil size (PCPS) (Iqbal et al., 2004).

A number of additional eye metrics were collected and calculated as well (see below). Their analysis helped further understand the specifics of students' reading behavior in each of the information resources.

- 5) Visit count (measured in counts) – the number of visits within an active AOI.

A visit was defined as the time interval between the first fixation on the active AOI and the end of the last fixation within the same active AOI.

- 6) Mouse click count (measured in counts) demonstrated how many times the participant left-clicked with the mouse on the active AOI. It was assumed that left mouse clicks were associated with clicks on hyperlinks embedded in the text and not just random clicking on the content. By analyzing this metric I wanted to explore whether or not and how frequently the participants used suggested hyperlinks.

## **Preparing Data for Analysis**

**Step 1. Defining areas of interest.** The following actions took place in order to prepare collected eye metrics for analysis. Each information resource was divided into several AOIs based on observations of students' information behavior and their overall interactions with information resources. Some of these AOIs served the primary informational purpose (e.g., main content) and some were supplementary (e.g., navigation menu). Defining several AOIs within each information resource was instrumental as it allowed to keep eye-movement data from different AOIs separate. Below are examples and description of AOIs defined in *UpToDate*, *Access Medicine*, and *Wikipedia* (Figures 8-10).

## AOIs in *UpToDate*

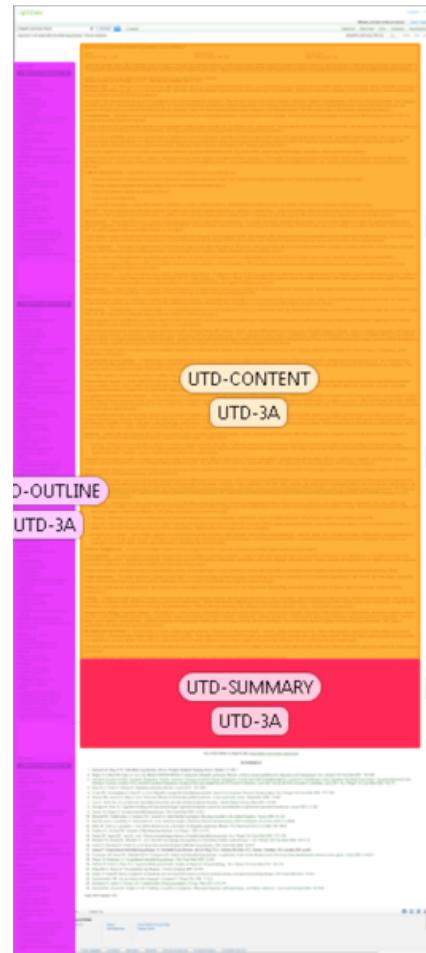


Figure 8. AOIs in *UpToDate*

- 1) CONTENT – the main content.
- 2) OUTLINE – the left-side navigation bar.
- 3) SUMMARY – the summary of the article located right before the references.

Summary was defined as a separate AOI because observations of students' behavior showed that some students tended to immediately scroll down the page and read the summary first before attending to the main content.

4) MEDIA – all media elements (e.g., images, charts, tables) accessible via hyperlinks embedded in the main content.

AOIs in *Access Medicine*

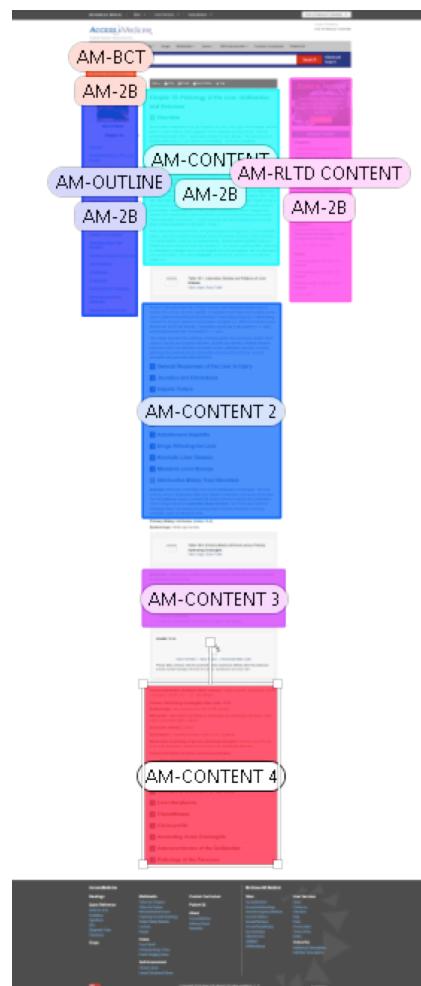


Figure 9. AOIs in *Access Medicine*

- 1) CONTENT – the main content.
- 2) OUTLINE – table of contents.
- 3) RELATED CONTENT – hyperlinks to related content in the form of chapters and multimedia.

4) MEDIA – all media (e.g., images, charts, tables) accessible on the page.

Because media in *Access Medicine* loaded as the user scrolled down the page, dynamic AOIs were created.

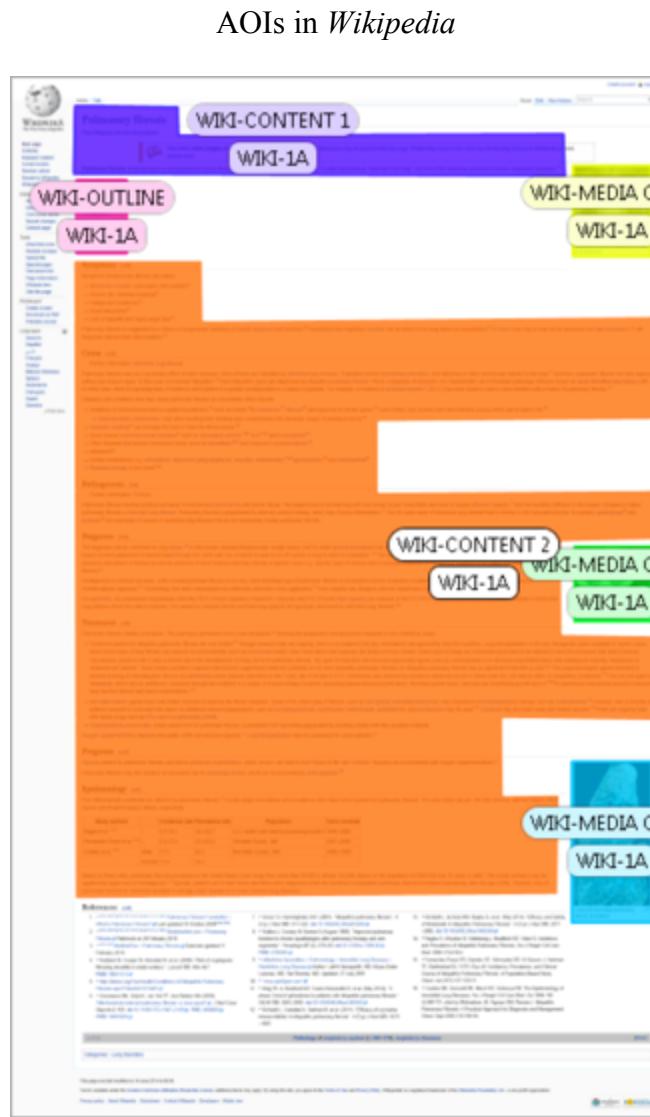


Figure 10. AOIs in *Wikipedia*

- 1) CONTENT – the main content.
- 2) OUTLINE – table of contents.
- 3) RELATED CONTENT – additional information about the topic.

4) MEDIA (on the right, on the left, in the center of the page, and in a separate window) – all media (e.g., images, charts, tables) accessible on the page or via hyperlinks.

Because the participants were instructed to remain strictly within the assigned information resource during each task, they were not allowed to click on references at the bottom of each content page; that is why, references to external resources were excluded from AOIs. The search box in all three information resources was also excluded from AOIs because observations revealed no instances of students performing searches – every new search was initiated on the main search page.

**Step 2. Computing eye metrics.** Based on the defined AOIs, Tobii studio yielded automatic calculations of all eye metrics of interest except for TEPR because it was a compound metric consisting of MPD, lability, and PCPS.

To calculate MPD, I needed to compute the difference between the values of pupil diameter over the duration of the task and the baseline pupil diameter (Beatty & Lucero-Wagoner, 2000). Several approaches have been reported in the literature regarding calculating the baseline pupil diameter. Metalis, Rhoades, Hess, & Petrovich (1980) averaged the pupil size over the ten seconds preceding the onset of the first stimulus and over the ten seconds following the offset of the stimulus, which resulted in pre-task and post-task baseline values. Iqbal et al. (2004) collected pupil baseline by having the user fixate on a blank screen for ten seconds before the user was presented with the actual task. Another group of researchers suggested establishing a pre-stimulus baseline period by averaging pupil data points during the first 500-1000 milliseconds (~ 0.5-1 seconds) of the task (Beatty & Lucero-Wagoner, 2000; Partala & Surakka, 2003).

To establish an accurate baseline period in this study, I reviewed all individual recordings and approximated that the relevant pre-stimulus interval was ~ 1000 msec. This was the time participants spent looking at the computer screen before the URL address of the stimulus was fully loaded and the content appeared on the screen. As soon as the content appeared on the screen, participants immediately attended to reading, which suggested the initiation of information processing. Data points during the pre-stimulus interval were averaged and used as pupil baseline measure in further calculations. Because each participant interacted with three information resources, I calculated three pupil baseline values per participant – one in each of the three information resources. Calculating pupil baseline relevant to each information resource was also considered appropriate since interface design characteristics of each information resource (e.g., background color) could have affected pupil diameter during pre-stimulus interval.

Lability was calculated as the difference between the maximum and the minimum values of pupil size during the task (Matalis et al., 1980). Both lability and MPD were reported in millimeters as a single value – an average of both eyes – for each information resource.

PCPS was calculated by detracting the average baseline pupil size from the average pupil size on the task and dividing it by the average pupil baseline (Iqbal et al., 2004; Matalis et al., 1980):

$$M\%X = [(MX-B) \times 100]/B$$

where  $M\%X$  - mean percentage dilation on task X,  $MX$  – mean pupil size during task X, and  $B$  – baseline pupil on task X.

PCPS is an informative metric; however, being a single value associated with the entire task, it does not allow to see the fluctuations in pupil size that took place during different stages of the task. That is why, in addition to PCPS values, I relied on visual representation of pupil fluctuations throughout the task. These visualizations helped identify moments of peak pupil dilation and link them to examples of peaks in cognitive processes.

### **Data Analysis for Research Question 1**

The following analysis was undertaken to answer the first research question:  
What are the differences in learners' mental effort invested in interactions with electronic information resources?

The differences in mental effort invested in interactions with electronic information resources were explored through the comparative analysis of total fixation duration, fixation count, total visit duration, and TERP. In this study, meaningful information acquisition occurred when the participants were reading the text; that is why, of primary interest were the comparisons of eye metrics on the CONTENT AOIs of the three information resources.

Repeated LSD analysis with observations nested under subjects was performed using a mixed model approach in IBM SPSS Statistics 19 software. Total fixation duration, fixation count, total visit duration, and TEPR were entered in the model as dependent variables. Results were expected to be significant at  $\rho \leq .05$ .

### **Data analysis for Research Question 2**

The following investigation was conducted to answer the second research question: How do the differences in learners' mental effort invested in interactions with electronic information resources affect learning?

Repeated LSD analysis with observations nested under subjects was performed using mixed model approach in IBM SPSS Statistics 19 software. Values of individual test results were entered in the model as a dependent variable and compared against the values associated with the mental effort, i.e., total fixation duration, fixation count, total visit duration, and TEPR. Results were expected to be significant at  $\rho \leq .05$ .

To help answer the second research question, I also investigated the effect of students' familiarity with diseases on the learning outcomes and their information behavior. Specifically, I hypothesized that those students who reported better familiarity and higher level of medical knowledge of the diseases would achieve higher scores and visit fewer unique pages per disease.

**Group comparison.** Because the knowledge test consisted of only ten questions, concerns were raised as to the power and validity of the tests. To address these concerns, an additional independent *t*-test was conducted to compare the test results of 18 participants of the main study with the test results of another nine subjects from the same population who took the test without preliminary information acquisition from information resources. The purpose of such comparison was to demonstrate that 1) prior knowledge of the diseases alone was not sufficient to achieve high scores on the knowledge tests; 2) electronic information resource played a significant role in knowledge acquisition; and 3) knowledge tests did measure learning and could not be

successfully completed without prior acquisition of relevant information from the evaluated information resources.

The additional nine subjects were recruited via mass e-mail. They consented to participate in the study and were instructed to take the electronic version of the same knowledge tests on the same three diseases. Each participant received \$25 for participation. Students were allowed to take the tests remotely, at the time and location convenient to them. Group 2 participants were encouraged to answer “I don’t know” instead of providing their best educated guess if they were not sure or did not know the answer to the test questions.

Because only nine students participated in this part of the experiment, only three Latin squares were randomly generated and assigned to participants (Table 6).

Table 6

*Three Latin Squares Used with Group 2*

Latin square #1

Participants	Diseases		
	IPF	NS	PSC
P20	<i>Wikipedia</i>	<i>UpToDate</i>	<i>Access Medicine</i>
P23	<i>Access Medicine</i>	<i>Wikipedia</i>	<i>UpToDate</i>
P28	<i>UpToDate</i>	<i>Access Medicine</i>	<i>Wikipedia</i>

Latin square #2

Participants	Diseases		
	PSC	IPF	NS

---

P21	<i>UpToDate</i>	<i>Access Medicine</i>	<i>Wikipedia</i>
P24	<i>Wikipedia</i>	<i>UpToDate</i>	<i>Access Medicine</i>
P25	<i>Access Medicine</i>	<i>Wikipedia</i>	<i>UpToDate</i>

---

Latin square #3

Participants	Diseases		
	NS	PSC	IPF
P22	<i>Access Medicine</i>	<i>Wikipedia</i>	<i>UpToDate</i>
P26	<i>Wikipedia</i>	<i>UpToDate</i>	<i>Access Medicine</i>
P27	<i>UpToDate</i>	<i>Access Medicine</i>	<i>Wikipedia</i>

---

### **Data Analysis for Research Question 3**

The following analysis was undertaken to answer the third research question:  
How do certain interface design characteristics of electronic information resources contribute to or interfere with information acquisition and, consequently, learning?

First, I wanted to understand whether availability and accessibility of hyperlinks had any effect on students' information behavior. For this purpose I compared mouse click counts on the CONTENT AOI of *UpToDate*, *Access Medicine*, and *Wikipedia*. Extensive use of hyperlinks, in its turn, could have resulted in differences in the number of unique pages visited by the students, which was also analyzed.

Second, analysis of interfaces revealed that media in *UpToDate* could be accessed only via hyperlinks while in *Access Medicine* and *Wikipedia* media were immediately

accessible as they were embedded in the content page. Therefore, I wanted to understand whether the differences in access to media in three electronic information resources could affect students' information behavior. I explored this question by calculating the number of unique media viewed or interacted with in each of the information resources.

Of interest were also differences in placement of media on the page, e.g., in *Access Medicine* media were placed in the center on the page, while in *Wikipedia* media were placed on the left, on the right, and in the center of the page. Calculating total visit duration on the left, right, and center media in *Wikipedia* was informative in understanding which media attracted more attention.

Third, I wanted to explore the pattern of interaction with the search pages (available in *UpToDate* and *Access Medicine*) and navigational elements (available in all three information resources) by comparing the amount of time (total visit duration) the participants spent on those elements. Analysis of eye metrics also allowed me to identify interface design elements that were not interacted with at all.

Finally, I was interested in exploring opportunities for improvement of information resource interfaces based on observed user behavior and user feedback. For this reason, interface designs of the three electronic information resources were analyzed from the standpoint of the principles of instructional design in hypermedia learning and a number of principles of human-computer interaction. User feedback was transcribed verbatim, reviewed several times, and coded for the commonly occurring themes. Observations were documented and reviewed several times in order to identify emergent themes.

## **CHAPTER 4 – RESULTS**

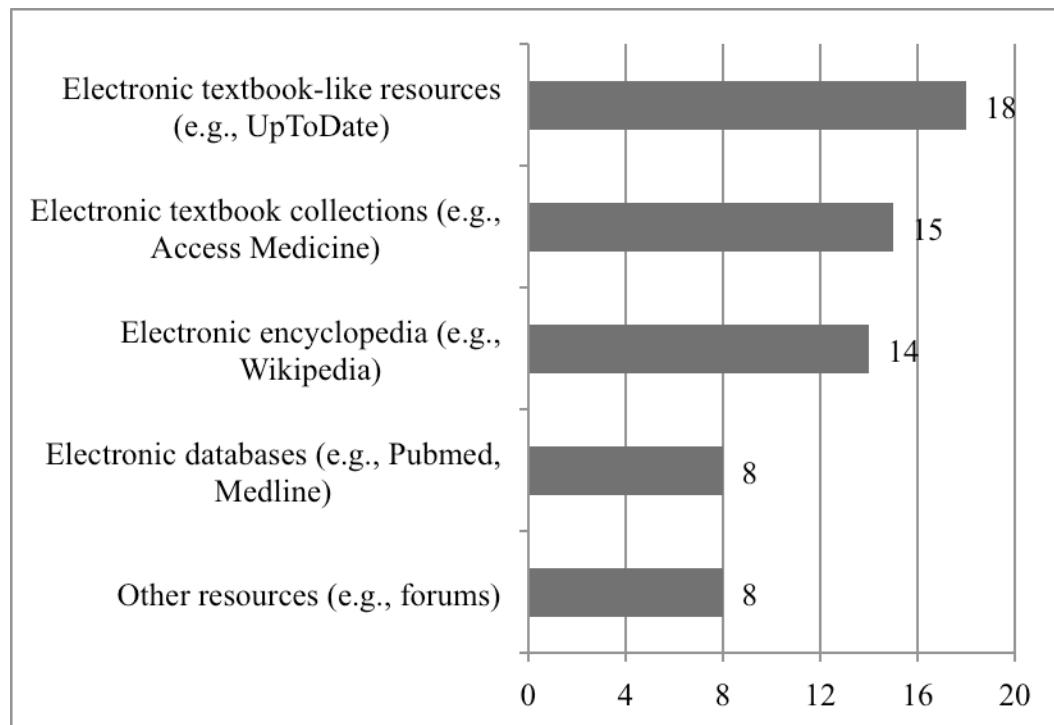
The chapter includes the results of data analysis aimed to reveal the differences in mental effort invested by learners in interactions with three electronic information resources of interest – *UpToDate*, *Access Medicine*, and *Wikipedia*. This chapter also includes the results of data analysis investigating the impact of mental effort on learning. Finally, this chapter presents an analytic review of interface design characteristics and functionality features of the electronic information resources and their role in contributing to or interfering with information acquisition and learning. Presentation of findings begins with the summary of participants' demographic information and overview of their preferences for electronic information resources.

### **Participants' Demographics**

The sample of participants included eleven male students and seven female first-year medical students of the MU SOM. Fifteen participants were in the age group between 22-26 years old and held bachelor's degrees (BSc=9 and BA=6). Three students were older than 27 years old and in addition to bachelor's degrees held master's degrees (MSc=2 and MA=1). Five students indicated that English was their second language.

### **Survey Results**

Participants indicated using a number of electronic information resources for learning (Figure 11).



*Figure 11.* Students' reported preferences for electronic information resources

Among other resources used for learning, students primarily mentioned hard copy textbooks and *Medscape* – a web resource for physicians and health professionals that features peer-reviewed original medical journal articles, daily medical news, a customized version of the National Library of medicine's MEDLINE database, and drug information ([www.medscape.com](http://www.medscape.com)). Students' answers to questions of why, what for, and how often they used each of these resources were analyzed and summarized in Table 7.

Table 7

*Reported Purpose, Frequency, and Reasons for Using Electronic Information Resources*

Information resource	Purpose of use	Frequency of use	Reasons for use
Electronic textbook collections (e.g.,	- Learning basic science - Learning disease processes	From daily to 3-4x/week	- Easy to understand - Easy to use

---

<i>Access Medicine</i> )	- Basic understanding of the subject		- Thorough, detailed - Credible - Recommended by faculty - Multiple resources to compare - Good summaries
Electronic textbook-like resources (e.g., <i>UpToDate</i> )	- Learning diseases processes (e.g., pathogenesis, diagnosis, management, treatment)  - Finding evidence  - Confirming hypothesis  - Clinical approach  - Learning basic science  - Diagnostic procedures  - General information  - Drug information	Daily	- Best compilation of primary sources  - More in-depth information  - Recommended by faculty  - Credible  - Easy to use  - Fast to read  - Easy to understand  - Used by faculty  - Up to date  - Good summaries  - Very thorough
Electronic encyclopedia (e.g., <i>Wikipedia</i> )	- Learning basic science  - Cursory review of everything  - Initial learning  - Quick reference for diseases  - Starting point for hypothesis/obtaining ideas	Daily	- Easy to understand  - Very fast to read  - Easy to understand  - Easy to access  - Concise

- Simple definitions
- Getting a basic broad understanding
- Reminder of concepts
- Good for things I don't know anything about
- To check for references and find other resources
- Helpful to refine search and knowing what to search for
- Finding evidence

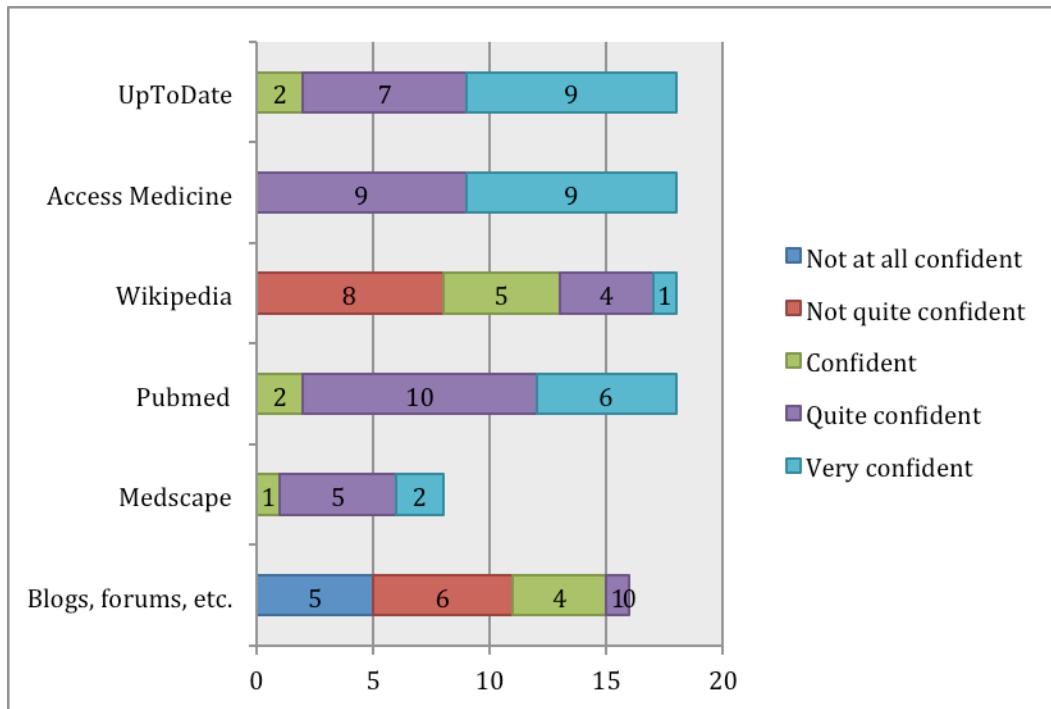
Electronic databases (e.g., Pubmed, MEDLINE)	<ul style="list-style-type: none"> <li>- Learning basic science (e.g., pathophysiology)</li> <li>- Finding the most recent research</li> <li>- Finding primary research</li> <li>- Disease review articles</li> <li>- Finding specifics about diseases</li> <li>- Finding evidence (e.g., treatment options)</li> </ul>	<p>From daily to weekly to occasionally</p>	<ul style="list-style-type: none"> <li>- Most recent information</li> <li>- Credible</li> <li>- Up to date</li> <li>- Recommended by faculty and library</li> <li>- Helpful for learning about rare diseases</li> </ul>
--	---	---	---

---

When asked about their confidence in the accuracy of information available in these electronic information resources, students on average reported similar levels of confidence on a 1 to 5 Likert scale (1 = not at all confident, 5 = very confident) for *Access Medicine* ( $M = 4.5$ ,  $SD = 0.51$ ), *UpToDate* ( $M = 4.39$ ,  $SD = 0.70$ ), *PubMed* ( $M = 4.22$ ,  $SD = 0.65$ ), and *Medscape* ( $M = 4.13$ ,  $SD = 0.64$ ). The average level of students'

confidence in the accuracy of information available in *Wikipedia* ( $M = 2.89$ ,  $SD = 0.96$ ) was as low as their confidence in the accuracy of information on the Internet in general (e.g., blogs, forums, websites) ( $M = 2.06$ ,  $SD = 0.93$ ).

Students' levels of confidence in the appropriateness of these electronic information resources for learning are presented in Figure 12 below.



*Figure 12.* Students' reported confidence in the appropriateness of electronic information resources for learning

### Research Question 1

What are the differences in learners' mental effort invested in interactions with different electronic information resources?

Before the Latin square design (LSD) analysis was performed, normality of dependent variables – total fixation duration, fixation count, total visit duration, and task

evoked pupillary response (TEPR) – was ascertained by running Kolmogorov-Smirnov and Shapiro-Wilk tests (Table 8).

Table 8

*Normality Test of Eye Metrics Associated with Mental Effort*

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Total_Fix_Dur	.071	54	.200	.969	54	.166
Fix_Count	.102	54	.200	.963	54	.092
Total_Visit_Dur	.088	54	.200	.959	54	.060
MPD	.140	54	.010**	.944	54	.014
LABILITY	.081	54	.200	.983	54	.618
PCPS	.081	54	.200	.962	54	.083

\*\*  $p < .01$

Mean pupil dilation (MPD) was the only variable that was not normally distributed; therefore, it was subjected to Log10 and Square root transformations. Transformations, however, did not improve the normality of this variable. As a solution, I analyzed z-values of skeweness of 0.618 ( $SE = 0.325$ ) and kurtosis of 1.759 ( $SE = 0.639$ ), which fell between  $\pm 1.96$ . This suggested that although data were not perfectly normally distributed, they could still be considered normally distributed.

Results of the repeated LSD revealed that there were statistically significant differences in total fixation duration, fixation count, and total visit duration of the learners when they interacted with three electronic information resources ( $F(2, 32) = 7.364, \rho = .002$ ). Post hoc Fischer's Least Significant Difference test indicated that *UpToDate* was associated with significantly higher values of total fixation duration, fixation count, and total visit duration than *Access Medicine* or *Wikipedia* (Table 9).

Table 9

*Significance of Eye Metrics Associated with Mental Effort*

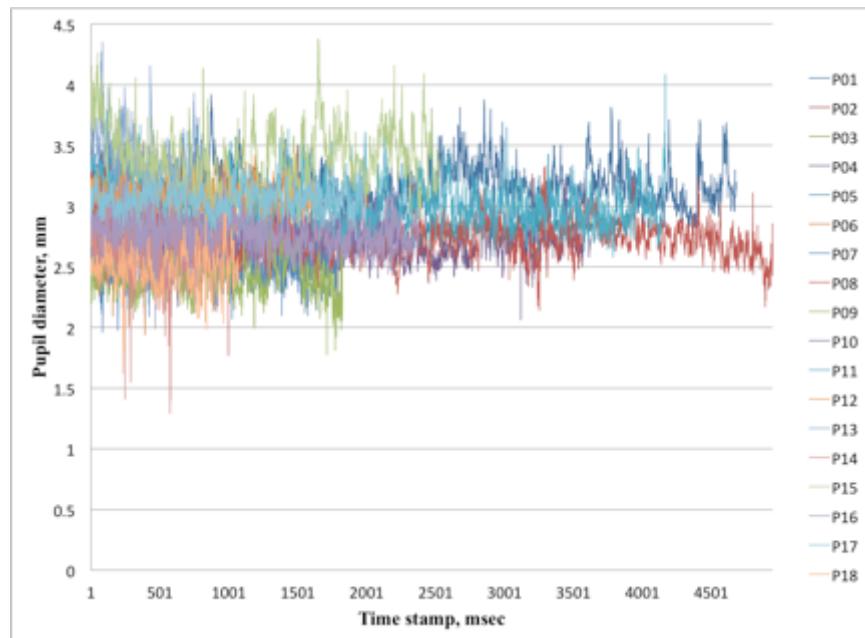
Parameter measured	<i>UpToDate</i> (mean $\pm$ SD)	<i>Access Medicine</i> (mean $\pm$ SD)	<i>Wikipedia</i> (mean $\pm$ SD)
Total fixation duration	464.22 $\pm$ 168.63**	302.81 $\pm$ 164.73	346.74 $\pm$ 172.58
Fixation count	2586.28 $\pm$ 1330.76**	1725.94 $\pm$ 856.39	1898.61 $\pm$ 869.21
Total visit duration	769.64 $\pm$ 344.67**	511.89 $\pm$ 232.78	524.05 $\pm$ 220.65

\*\*  $\rho < .01$ , indicates significance from *Access Medicine* and *Wikipedia*

Results of the repeated LSD analysis did not reveal statistically significant difference between learners' TEPR when interacting with three electronic information resources ( $\rho > .05$ ). Nonetheless, a number of differences in measures contributing to TEPR were found through descriptive statistics. For example, participants' MPD, on average, was 2.86 mm during the tasks performed both in *UpToDate* ( $M = 2.86, SD = 0.30$ ) and *Wikipedia* ( $M = 2.86, SD = 0.31$ ) and slightly wider during task performance in

*Access Medicine* at 2.90 mm ( $M = 2.90$ ,  $SD = 0.48$ ). Lability values showed that, on average, the difference between maximum and minimum pupil size throughout the tasks was 1.40 mm both in *UpToDate* and *Access Medicine* and 1.49 mm in *Wikipedia*. Values of PCPS showed that during information acquisition from three information resources participants' pupil tended to constrict from baseline. To be exact, participants' pupil on average constricted by 9.46 % (or 0.33 mm) during information acquisition from *UpToDate*, by 7.41 % (or 0.23 mm) from *Access Medicine*, and by 6.56 % (or 0.22 mm) from *Wikipedia*.

Visualization of pupillary activity of participants throughout the task per information resource allowed to see fluctuations in pupil diameter (Figures 13-15).



*Figure 13. Pupil diameter fluctuations among participants using UpToDate*

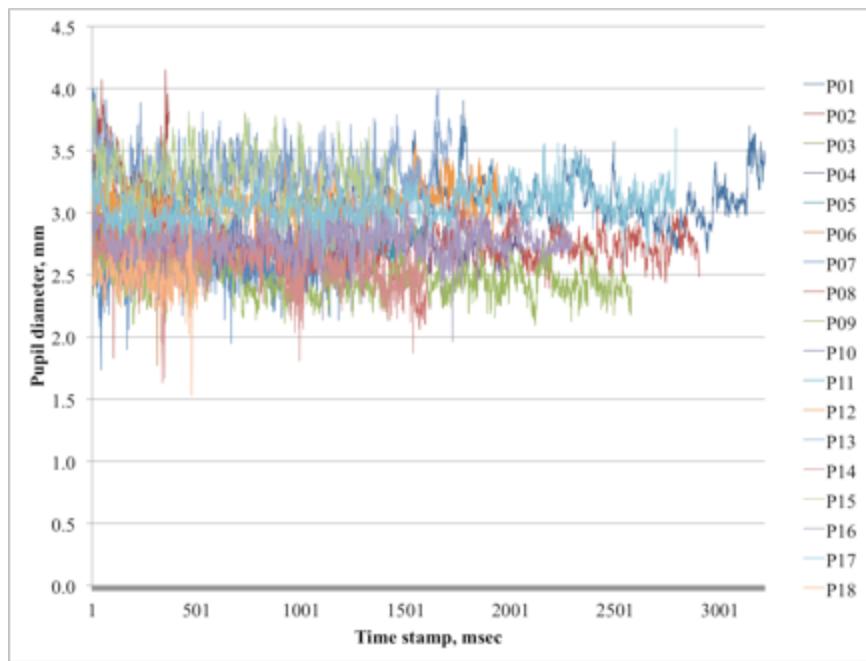


Figure 14. Pupil diameter fluctuations among participants using *Access Medicine*

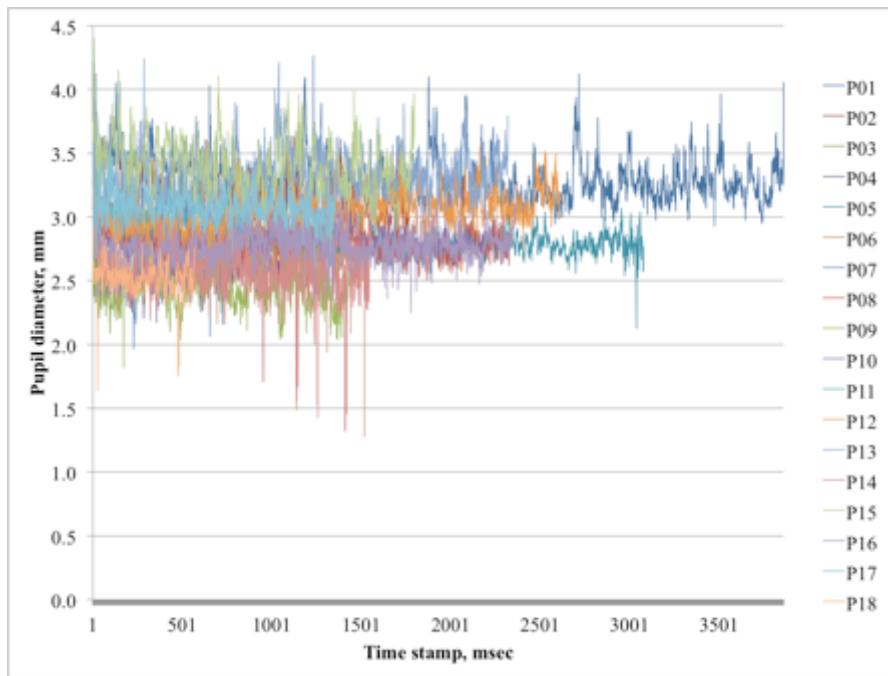


Figure 15. Pupil diameter fluctuations among participants using *Wikipedia*

Additional review of the recordings was done to pinpoint the exact moments during the task that corresponded to spikes of pupillary activity as seen on the charts. These spikes usually corresponded to longer fixations and coincided with fixating on unfamiliar words (Figure 16).

**Original Version (Top):**

earliest event in generation of the noncaseating granuloma. The CD4<sup>+</sup> T cells result from oligoclonal expansion, suggesting development of a conventional antigen-stimulated immune response.

The variation in incidence, severity, and manifestations of disease among different racial and ethnic groups suggests a genetic predisposition to disease development. Another observation that supports genetic susceptibility is familial clustering of disease. For example, sarcoidosis occurs two to four times more frequently in monozygotic than in dizygotic twins. Similarly, up to 19% of affected African-American families and 5% of white families have more than one affected family member. Results from a case-control etiological study of sarcoidosis (ACCESS) further support these observations. This study enrolled over 700 case-control pairs in the United States matched for age, gender, race, and ethnicity. Analysis of nearly 11,000 first-degree and over 17,000 second-degree relatives of these groups demonstrated an overall adjusted familial relative risk of developing sarcoidosis of 4.7 (95% CI = 2.3–9.7). In this study whites had a much higher familial relative risk compared to African-Americans (18.0 versus 2.8;  $p < 0.001$ ). It is likely given these observations that multiple genes, rather than a single gene, are involved in the genetic predisposition to sarcoidosis.

Studies of polymorphisms in the HLA family of genes have yielded inconsistent results. Associations between specific HLA-DR, -DQ, and -DP alleles and the presence of sarcoidosis or specific characteristics of the disease have been identified, and the associated alleles have differed based on race and ethnicity. In contrast, the ACCESS study results demonstrate a significant association between sarcoidosis and HLA-DRB1\*1101 across the entire patient cohort, including African-Americans and whites. Despite these and other associations of genetic polymorphisms with disease, the exact nature of the genetic predisposition to sarcoidosis remains unclear.

Rossman M et al. Association with human leukocyte antigen class II amino acid epitopes and interaction with environmental exposures. *Chest* 2002;121:184. [PubMed: 11893656]

**Redacted Version (Bottom):**

earliest event in generation of the noncaseating granuloma. The CD4<sup>+</sup> T cells result from oligoclonal expansion, suggesting development of a conventional antigen-stimulated immune response.

The variation in incidence, severity, and manifestations of disease among different racial and ethnic groups suggests a genetic predisposition to disease development. For example, sarcoidosis occurs two to four times more frequently in monozygotic than in dizygotic twins. Similarly, up to 19% of affected African-American families and 5% of white families have more than one affected family member. Results from a case-control etiological study of sarcoidosis (ACCESS) further support these observations. This study enrolled over 700 case-control pairs in the United States matched for age, gender, race, and ethnicity. Analysis of nearly 11,000 first-degree and over 17,000 second-degree relatives of these groups demonstrated an overall adjusted familial relative risk of developing sarcoidosis of 4.7 (95% CI = 2.3–9.7). In this study whites had a much higher familial relative risk compared to African-Americans (18.0 versus 2.8;  $p < 0.001$ ). It is likely given these observations that multiple genes, rather than a single gene, are involved in the genetic predisposition to sarcoidosis.

Studies of polymorphisms in the HLA family of genes have yielded inconsistent results. Associations between specific HLA-DR, -DQ, and -DP alleles and the presence of sarcoidosis or specific characteristics of the disease have been identified, and the associated alleles have differed based on race and ethnicity. In contrast, the ACCESS study results demonstrate a significant association between sarcoidosis and HLA-DRB1\*1101 across the entire patient cohort, including African-Americans and whites. Despite these and other associations of genetic polymorphisms with disease, the exact nature of the genetic predisposition to sarcoidosis remains unclear.

Rossman M et al. Association with human leukocyte antigen class II amino acid epitopes and interaction with environmental exposures. *Chest* 2002;121:184. [PubMed: 11893656]

Figure 16. Example of a longer fixation duration

## **Research Question 2**

How do the differences in learners' mental effort invested in interactions with electronic information resources affect learning?

In this study, learning was evaluated with the help of three tests, each containing ten multiple-choice questions pertinent to each of the three diseases. In preparation for normality check of the test results, two transformations were performed. First, the proportion of correct answers was calculated by dividing each test score by ten. Then, Arcsine transformation was performed in order to help make the distribution of data on proportions normal. Normality of the transformed test variable was ascertained through Kolmogorov-Smirnov and Shapiro-Wilk tests (Table 10).

Table 10

*Normality Test of the Knowledge Test Variable*

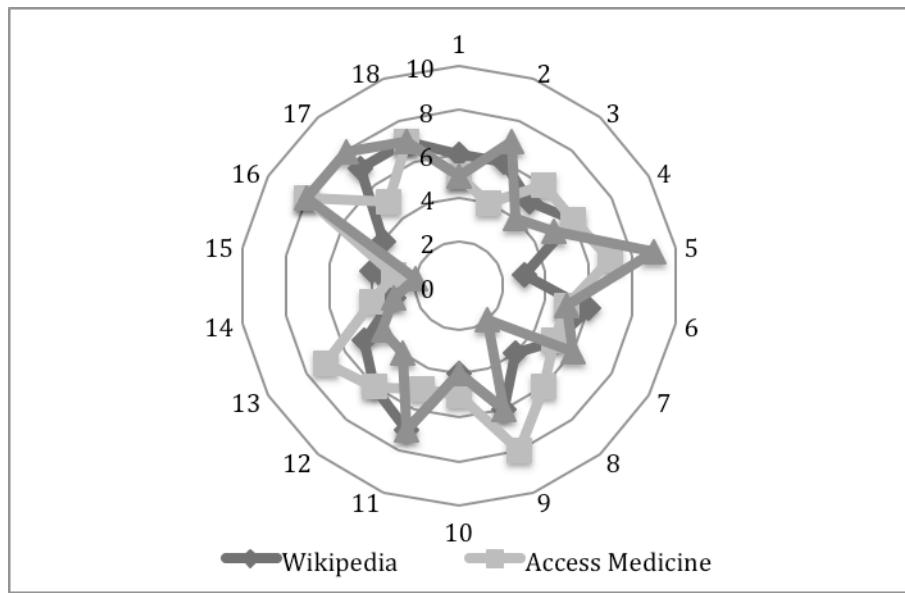
	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
TEST_Arcsin	.129	54	.026*	.966	54	.128

\*  $\rho < .05$

Shapiro-Wilk test was not significant ( $\rho = .128$ ); however, Kolmogorov-Smirnov was ( $\rho = .026$ ). This suggested that the distribution of the sample was significantly different from normal, i.e., not normal. To address this discrepancy, I plotted data on normal Q-Q plot to visually explore it and analyzed z-values of skeweness and kurtosis. The Q-Q plot showed that data were not normally distributed. Z-value for skeweness of 0.014 ( $SE = 0.325$ ) and kurtosis of -0.134 ( $SE = 0.639$ ), however, fell between  $\pm 1.96$ ,

which suggested that although data were not perfectly normally distributed, it could be considered normally distributed.

Results of the repeated LSD demonstrated no statistically significant difference between the values of learner's mental effort (as measured through total fixation duration, fixation count, total visit duration, and TEPR) and learning ( $\rho > .05$ ). However, of interest were the results of descriptive statistical analysis of test scores. They demonstrated that on average participants were able to answer correctly five out of ten questions when using *UpToDate* ( $M = 5.33$ ,  $SD = 2.06$ ), *Access Medicine* ( $M = 5.72$ ,  $SD = 1.36$ ), or *Wikipedia* ( $M = 5.22$ ,  $SD = 1.31$ ) (Figure 17).



*Figure 17. Distribution of knowledge test scores*

Participants were also able to answer five out of ten questions on average per disease, e.g., IPF ( $M = 5.06$ ,  $SD = 1.80$ ), PSC ( $M = 5.61$ ,  $SD = 1.58$ ), and NS ( $M = 5.61$ ,  $SD = 1.42$ ). Further analysis of students' responses demonstrated that the highest number of "I don't know"/"I can't remember reading about it" answers was given after students interacted with *Access Medicine* (n=29) compared to *UpToDate* (n=28) and *Wikipedia*.

(n=22). The highest number of “I don’t know”/“I can’t remember reading about it” answers was also given to the test questions on IPF (n=37) compared to PSC (n=22) and NS (n=20). These findings suggested that on average students reported not knowing answers two out of ten times when answering IPF questions ( $M = 2.06$ ,  $SD = 1.47$ ) and when using *Access Medicine* ( $M = 1.61$ ,  $SD = 1.2$ ).

Participants reported different degree of prior knowledge and familiarity with the diseases. Eight students reported being familiar to a certain degree with PSC and five with IPF. Students’ familiarity with these diseases varied and originated primarily from hearing about the disease as part of their coursework and/or brainstorming differential diagnoses on PBL cases. Further, despite being familiar with the diseases, students mainly reported having poor to no medical knowledge of these diseases (Table 11).

Table 11

*Reported Levels of Familiarity and Prior Knowledge of the Diseases*

	<b>IPF</b>	<b>PSC</b>	<b>NS</b>	<b>Total</b>
Familiarity with the diseases				
Not at all	13	10	16	39
A little	3	5	2	10
Somewhat	1	3	-	4
Familiar	1	-	-	1
Very familiar	-	-	-	-
Prior knowledge of the diseases				
None	12	11	17	40
Poor	5	7	1	13
Fair	-	-	-	-

Good	1	-	-	1
Excellent	-	-	-	-

Students' prior knowledge came from studying the signs and symptoms of diseases as part of their coursework, reading about the diseases when taking open book exams, or general knowledge of human physiology. In addition, none of the students reported being emotionally attached to any of the three diseases (e.g., none had friends or family members diagnosed with these diseases).

Results of the repeated LSD intended to explore whether different levels of students' familiarity with and prior knowledge of the diseases had any effect on learning were not significant ( $\rho > .05$ ). Despite this fact, the differences in students' prior knowledge of the diseases could potentially impact the amount of information they sought. Regardless of the disease, participants on average visited more unique pages when acquiring information from *Wikipedia* ( $M = 5, SD = 2.45$ ) than from *Access Medicine* ( $M = 4, SD = 1.31$ ) or *UpToDate* ( $M = 3, SD = 1.24$ ). Comparison of the number of unique pages visited by participants when learning about the same topic showed that students on average visited more unique pages when reading about PSC in *Wikipedia* (n=7) and more unique pages when reading about NS in *Access Medicine* (n=5). When reading about IPF, participants visited on average four pages when reading from *UpToDate*, *Access Medicine*, and *Wikipedia* (Table 12).

Table 5

*Average Number of Visited Unique Pages*

Disease	UpToDate	Access Medicine	Wikipedia	Total

IPF	4	4	4	12
PSC	3	3	7	13
NS	3	5	4	12
Total	10	12	15	

### Group Comparison

Comparison between the two groups was intended to demonstrate whether information acquired from electronic information resources contributed to learning and whether students' prior knowledge was sufficient for answering the test questions.

Group 2 was represented by six female and three male students. Seven students in the age group between 22-26 years old held bachelor's degrees (BSc=5 and BA=2) and two students in the age group of 27 and older held master's degrees (MA=1 or MSc=1). All nine students spoke English as their first language. Participants in Group 2 reported the following preferences for electronic information resources for learning (Figure 18).

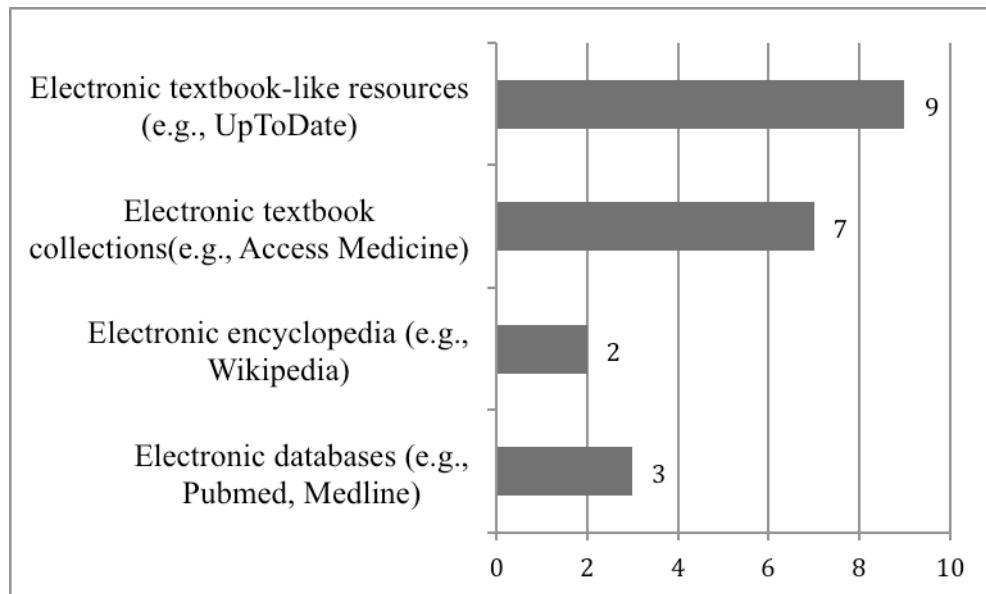


Figure 18. Group 2 preferences for electronic information resources for learning

Screening tests showed that students in Group 2 were most familiar with PSC (n=6) and IPF (n=5) and least familiar with NS (n=2). Six students had a certain degree of prior knowledge of IPF, three students of PSC, and only one student of NS. One student reported being emotionally attached to PSC because she encountered a patient with PSC during Advanced Clinical Examination.

Descriptive statistics of test results in Group 2 showed that, on average, participants received the highest score of 2 on the test on IPF ( $M = 2.11$ ,  $SD = 1.17$ ) and the lowest score of 0 on PSC ( $M = 0.56$ ,  $SD = 0.88$ ). The highest number of “I don’t know” answers was associated with NS (n=79) followed by PSC (n=72) and IPF (n=55). On average, participants in Group 2 did not know the answer to the test questions on NS almost nine (8.78) times.

In preparation for an independent samples  $t$ -test, test scores of the three tests for each participant in Groups 1 and 2 were averaged to obtain a single value. After normality was established, the results of the  $t$ -test revealed that, on average, participants in Group 1 performed significantly better on the test ( $M = 5.42$ ,  $SE = .27$ ) than participants in Group 2 ( $M = 1.11$ ,  $SE = .20$ ),  $t(25) = 10.463$ ,  $p = .000$ .

### **Research Question 3**

How do certain interface design characteristics of electronic information resources contribute to or interfere with information acquisition and, consequently, learning?

#### **Hyperlinks**

Descriptive statistical analysis of mouse click count demonstrated that throughout the task participants on average clicked within the main content AOI ten times in *UpToDate* ( $M = 10$ ,  $SD = 8.51$ ) and seven times in both *Access Medicine* ( $M = 7$ ,  $SD =$

6.77) and *Wikipedia* ( $M = 7$ ,  $SD = 5.38$ ). The number of mouse clicks, however, did not correspond to the number of the visited unique content pages in the three information resources. Therefore, it was assumed that some of the mouse clicks in *UpToDate* resulted from opening media because it was not immediately available from the content pages.

In *UpToDate*, hyperlinks were of two types: (1) one- or two-word hyperlinks and (2) sentence-long hyperlinks (Figure 19).

**Gallbladder cancer** — The prevalence of gallbladder cancer among patients with PSC is estimated to be 3 to 14 percent [72]. We agree with the American Association for the Study of Liver Diseases [3] and the European Association for the Study of the Liver [74], which recommend annual screening for gallbladder cancer using ultrasound. However, MRCP may permit screening for both gallbladder cancer and cholangiocarcinoma [72]. (See "[Gallbladder polyps and cholesterosis](#)", section on "[Polyps along with gallstones or primary sclerosing cholangitis](#)".)

For patients with mild-to-moderate IPF based on pulmonary function tests who live in an area where either [pirfenidone](#) or [nintedanib](#) is available, we recommend initiating therapy with the available agent (**Grade 1B**). Current data are insufficient to direct a firm choice between nintedanib and pirfenidone. Patient preference and tolerances should be considered, particularly regarding potential adverse effects. (See '[Our approach](#)' above and '[Assessing disease severity](#)' above.)

*Figure 19.* Types of hyperlinks in *UpToDate*

Both hyperlinks were in parenthesis, underscored, preceded by the word 'See', and in dark green color. One-word hyperlinks typically linked to information on drugs, while sentence-long hyperlinks linked to other thematic articles. Sentence-long hyperlinks usually appeared at the end of the sentence, which was perceived by some students as interfering with reading continuity. *UpToDate* hyperlinks that were sentence long, e.g., (See "[Clinical manifestations, diagnosis, and natural history of primary biliary cirrhosis](#)", section on '[Liver biopsy](#)') were broad in scope. Because of this, students had no way to quickly look up a definition or explanation of just an unfamiliar word or concept. Rather, in order to understand something in the first article, students had to read

the second article in full first. At the same time, there were a few participants who really liked this type of hyperlinks as they could read beyond the current topic.

During the interview, some students commented about the destination of the hyperlinks, i.e., sometimes following a hyperlink took them to a different spot within the same article, and sometimes – to a different article. Not knowing where exactly within the information resource students were after following a hyperlink made them feel like “*reading around in a circle.*” As one participant commented: “*In UpToDate, it’s kind of frustrating, if you click on it [hyperlink] and it [hyperlink] takes you to a whole other article and then you have to go back to the one you were reading before, which I guess is the same way as Wikipedia, but I don’t know why I just think Wikipedia is easier in that respect.*” Closer review of the hyperlinks format, however, showed that hyperlinks that linked to a different spot within the same page had indicators ‘below’ or ‘above’, e.g., (See ‘Prognosis’ below) or (See ‘Natural history’ above). It is possible that during quick reading, these indicators did not stand out to the participants; therefore, were overlooked.

*Access Medicine* had very few hyperlinks embedded in the main content. These hyperlinks typically took users to a different book or a different chapter of the same book, pages with media or information about drugs. There was nothing notable in students’ comments about or their use of hyperlinks in *Access Medicine*.

Hyperlinks in *Wikipedia* were highlighted in bright blue color, neither in parenthesis nor underscored, and consisted primarily of one or two words, e.g., ‘blood flow’, ‘anemia’. It was notable to observe that as students read the main content and encountered an unfamiliar word, they almost always right-clicked on it and opened it in a new browser tab. Some students continued to read by switching between the tabs; others

preferred to finish reading the main page before they started to read additional pages. On these additional pages, often students read only a few lines of the topic summary, available at the top of the page. This was reportedly sufficient for them to get an idea of what the new word meant and smoothly proceed with the primary topic (Figure 20).



*Figure 20. Example of a quick read from Wikipedia*

Students considered hyperlinks in *Wikipedia* a quick and easy way to obtain access to needed information. At the same time, they cautioned against their extensive use as not to get carried away too far from the main topic. As one student commented: “*I think with Wikipedia, a lot of the times you can get lost down like a rabbit trail, where you just keep clicking it [hyperlink] and keep going further and further, which is really cool. I like to do that. I think it helps you learn, but sometimes you get way outside of what you were actually looking at in the first place.*”

### **Media (i.e., images, simulations, audio recordings)**

Observations of students’ interactions with information resources showed that during the tasks participants viewed a different number of unique media per resource. On average, students viewed more media in *Access Medicine* ( $M = 3$ ,  $SD = 1.82$ ) compared

to *UpToDate* ( $M = 1$ ,  $SD = 1.23$ ) or *Wikipedia* ( $M = 2$ ,  $SD = 2.3$ ). Accordingly, participants spent more time (total visit duration in seconds) looking at the media in *Access Medicine* ( $M = 35.6$ ,  $SD = 29.17$ ) compared to *UpToDate* ( $M = 11.83$ ,  $SD = 16.12$ ) or *Wikipedia* ( $M = 15.25$ ,  $SD = 9.69$ ). Below I describe the difference in the placement and format of the media in the three information resources, which I believe that could have contributed to differences in the viewing behavior.

In *UpToDate*, the media were accessible only via hyperlinks embedded in the main content. Hyperlinks leading to media were typically termed ‘image 1’ or ‘table 2’, parenthesized, and appeared several times throughout the article. Unless the hyperlink leading to an image or a table was clicked, there was no way to tell if media would be helpful in further understanding of the topic. Absence of a thumbnail or a description of the image or table sometimes made it difficult for students to remember whether or not they had clicked on these media previously. As a result, students had to either retain in their short-term memory the information of which images or tables they had already clicked or continue to repetitively open the same media without the actual need to do so. To eliminate this frustration, some students reported completely avoiding using media in *UpToDate*. Those who did click on the hyperlinks leading to images reported liking the fact that images opened in a separate pop-up window, right in front of the window with the main content. Students also reported that reading the text under the images in addition to the image itself was helpful in understanding the concept (Figure 21).

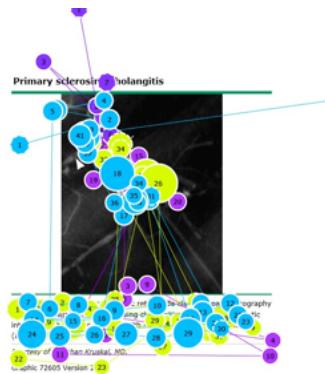


Figure 21. Example of viewing image in *UpToDate*

In *Access Medicine* media (e.g., histology and x-ray illustrations, diagrams) were referenced within the text and immediately available from the content pages (Figure 22). Some students perceived placement of media between the text blocks as interruption of textual flow, which provided a timely break from reading. Despite the fact that thumbnails of *Access Medicine* media embedded in text were found to be rather small, they were still reported helpful in allowing students to decide whether or not the media was of interest to them. The major complaint about the media on the page was their loading speed. According to students, the fact that media were well referenced within the text helped them integrate information from text and images, which promoted their understanding of the concepts.



*Figure 22. Layout of media in Access Medicine*

In *Wikipedia* most media were immediately available from the main content pages and arranged either in the center of the page or on both sides (Figure 23). Some students perceived immediate access to media as beneficial for learning. Students could quickly glance at the images and continue reading. These students thought they could grasp the meaning of the concept much faster by simply looking at the image than by reading textual description of the same concept.

Others, on the other hand, thought of permanent image placement in several parts of the page as an attempt to make them look at the images all the time and found it distracting, especially when the images were of certain deformed body parts, e.g., eyes (Graves disease, clubbing of fingernails). These participants also did not find it helpful when the same images, for example, depicting basic genetic processes, appeared over and

over again in several related articles. An interesting piece of feedback was that images in *Wikipedia* were not referenced in the text, which made them look disjointed to some students. To others, absence of references to images in the text made these images unnoticeable, especially if the images were on either side of the page.



*Figure 23. Layout of media in Wikipedia*

Comparison of participants' total viewing time of media in different locations on *Wikipedia* pages showed that during the experiment the participants on average spent more time looking at the media in the center of the page ( $\sim 7$  sec) ( $M = 6.96, SD = 15.25$ ) than those on the left ( $\sim 3$  sec) ( $M = 3, SD = 9.85$ ) or on the right ( $\sim 3$  sec) ( $M = 2.93, SD = 5.05$ ). Still, in *Wikipedia* students looked at the media placed in the center of the pages

about five times less than media placed in the center of the pages in *Access Medicine* ( $M = 35.6$ ,  $SD = 29.17$ ).

## Search Pages

Search pages were available only in *UpToDate* and *Access Medicine*. In *UpToDate*, the search page consisted of several sections (Figure 24). It contained a list of titles of the articles relevant to the search (colored in dark green). Each article was further subdivided into several sections (e.g., clinical manifestations, diagnosis, summary and recommendations) (colored in light blue). Each section, in its turn, had a topic outline as a separate extended menu that provided a more detailed view of the content. On the left of the search page, the user could filter the search results in accordance with five categories, i.e., all topic, adult, pediatrics, patient, and graphics.

The screenshot shows the UpToDate search results for the query "primary sclerosing cholangitis". The search bar at the top contains the query. Below it, a sidebar on the left allows filtering by category: All Topics (selected), Adult, Pediatric, Patient, and Graphics (with 5 items). The main content area displays a list of search results, each with a dark green title and a light blue outline of sub-sections. A large callout box highlights the "Topic Outline" for the first result, "Primary sclerosing cholangitis in adults: Clinical manifestations and diagnosis". This outline includes sections like "SUMMARY AND RECOMMENDATIONS", "INTRODUCTION", "NATURAL HISTORY", "ASSOCIATION WITH INFLAMMATORY BOWEL DISEASE", "CLINICAL MANIFESTATIONS", "DIAGNOSIS", "DIFFERENTIAL DIAGNOSIS", and "COMPLICATIONS OF PRIMARY SCLEROSING CHOLANGITIS". Each section has a list of bullet points. At the bottom of the page, there are links for "Languages" and "Help".

Figure 24. Search page in *UpToDate*

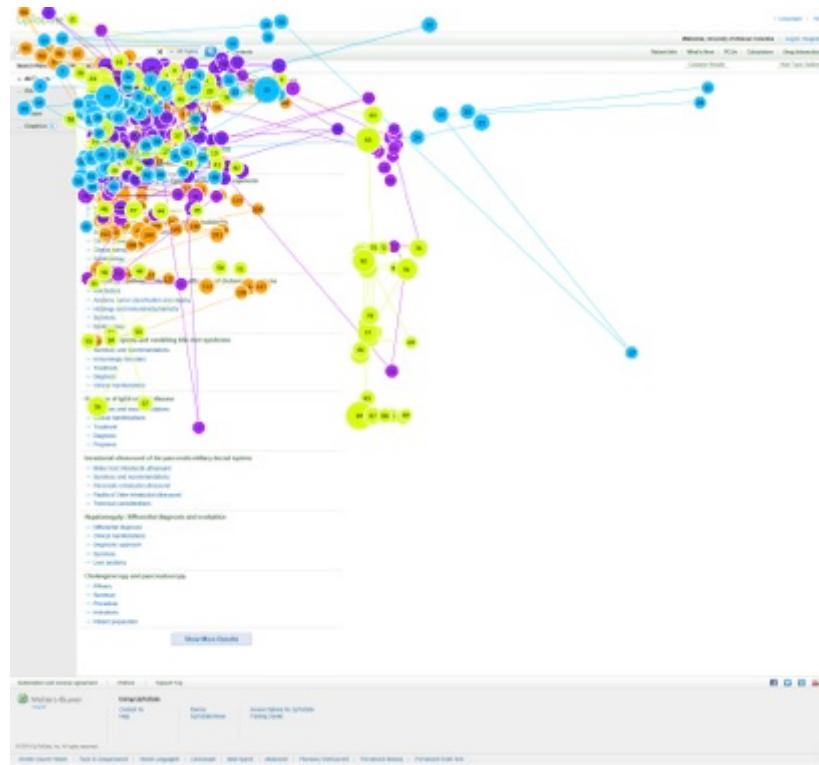
In *Access Medicine*, after the search was performed, the user was presented with a list of textbooks and textbook chapters (Figure 25). Users had the opportunity to narrow the search by topic or the type of content (i.e., reading, images, or cases). On the right side of the search page they also had access to relevant images and cases.

The screenshot shows the Access Medicine interface. At the top, there's a navigation bar with links for Home, Readings, Quick Reference, Drugs, Multimedia, Cases, Self-Assessment, Custom Curriculum, and Patient Ed. Below the navigation bar is a search bar with the placeholder 'Search AccessMedicine' and a red 'Search' button. To the right of the search bar is a link to 'Advanced Search'. The main content area has a heading 'primary sclerosing cholangitis' and a note: 'Search only displays content that you own, as long as you are signed in.' On the left, there's a sidebar with a 'Readings (100)' section containing 'Textbooks (100)' which is highlighted in blue, followed by 'Updates (11)', 'Quick Dx and Rx (0)', 'Diagnosaurus (7)', 'Diagnostic Tests (2)', 'Images (13)', and 'Cases (2)'. The main content area has a 'Textbooks' section with a 'Narrow By Topic' link and a note '160 results in Textbooks'. It lists several textbook chapters: 'Chapter 52. Primary Sclerosing Cholangitis' from CURRENT Diagnosis & Treatment: Gastroenterology, Hepatology, & Endoscopy, 2e; 'Chapter 15. Pathology of the Liver, Gallbladder, and Pancreas > Primary Sclerosing Cholangitis (See Table 15-6)' from Pathology: The Big Picture; 'Chapter 11. Bile Formation and Secretion > Primary Biliary Cirrhosis and Primary Sclerosing Cholangitis' from Gastrointestinal Physiology, 2e; 'Chapter 308. Cirrhosis and Its Complications > Primary Sclerosing Cholangitis' from Harrison's Online; and 'Chapter 308. Cirrhosis and Its Complications > Treatment: Primary Sclerosing Cholangitis' from Harrison's Online. On the right, there's an 'Images' section with three items: 'Figure e38-23 Primary sclerosing cholangitis showing cirrhosis and periductular fibrosis (Masson trichrome, 4x)...' (Harrison's Online > Chapter e38, Atlas of Liver Biopsies), 'Figure e38-24 Primary sclerosing cholangitis showing the extrahepatic bile duct (in a liver explant obtained at the time of hepatectomy)....' (Harrison's Online > Chapter e38, Atlas of Liver Biopsies), and 'Figure e38-25 Primary sclerosing cholangitis showing peripheral cholestasis (green) and cytoplasmic red granular staining of hepatocytes for copper....' (Harrison's Online > Chapter e38, Atlas of Liver Biopsies). Below these images is a link 'All Images Results (13) >'. At the bottom of the right sidebar, there's a 'Cases' section with a 'CASE FILES' link and a 'Crohn Disease Case File' link.

*Figure 25. Search page in Access Medicine*

Comparison of total visit duration on the search pages showed that during the task students on average spent about 24 seconds on the SEARCH AOI in *UpToDate* ( $M = 24.10$ ,  $SD = 18.49$ ) and almost twice as long (i.e., 44.4 seconds) on the SEARCH AOI in *Access Medicine* ( $M = 44.40$ ,  $SD = 19.06$ ). Students' feedback from the interviews and overview of the interface design features of both information resources provided a few opportunities for understanding the differences in the observed behavior. First, due to

specific content listed under each article title, *UpToDate* search page could have been perceived as more detailed and, ultimately, more helpful in allowing to decide which articles or their sections to select (Figure 26).



*Figure 26.* Example of viewing search page in *UpToDate*

*Access Medicine*, on the other hand, only provided the image of a book and the title of the chapter that contained the sought information, which were the only criteria on which students based their decisions with regards to selecting the search results. Often, students had to check several search results before they found the one that best met their information needs. One student commented: “*I found Access Medicine to be the hardest resource to use because the information was all over the place and organized by textbook instead of by topic. If I needed to know about this [topic], I couldn't just click on it and go there, I would have to go back and find another textbook on that section.*” Review of

the search page interface design, however, showed that despite student complaints about not being able to search by the topic, this option was available –‘Narrow By Topic’ and only one student actually used it during the experiment (Figure 27).



*Figure 27. Example of viewing search page in Access Medicine*

## Navigation

*UpToDate* the navigation menu was placed to the left from the main content and permanently remained there throughout the whole length of the page (Figure 28).

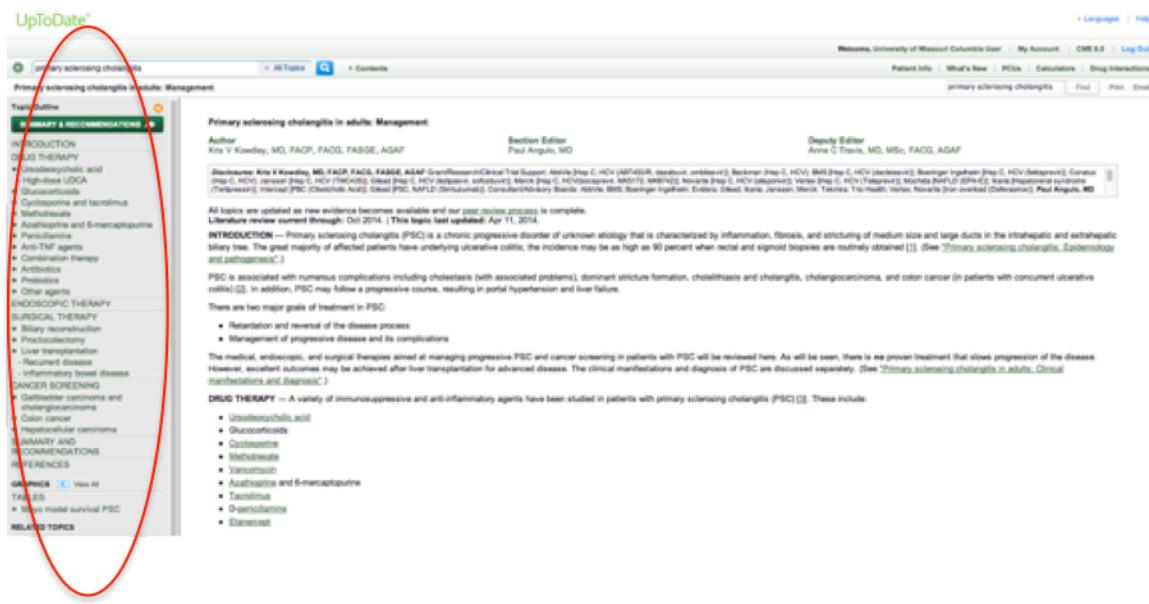


Figure 28. Left-side navigation menu in *UpToDate*

Throughout the task duration (~ 20 min), participants on average looked at the left navigation menu in *UpToDate* about 24 times ( $M = 24.05$ ,  $SD = 40.39$ ) totaling in 13.5 seconds ( $M = 13.5$ ,  $SD = 21.92$ ) but clicked on it only twice ( $M = 2.16$ ,  $SD = 3.05$ ).

Navigation in *Access Medicine* was represented by several elements (Figure 29). First, the outline of the book or a chapter was located on the left side of the page under the image of the book. Compared to the outline in *UpToDate*, it was less detailed and much shorter. Second, in some textbooks, additional related content was available to the right from the main content. This related content included references to other chapters of the same book that discussed the same topic, information about textbook updates, and

cases on the same topic. Third, the breadcrumb navigation trail was available in the top left corner of the page.

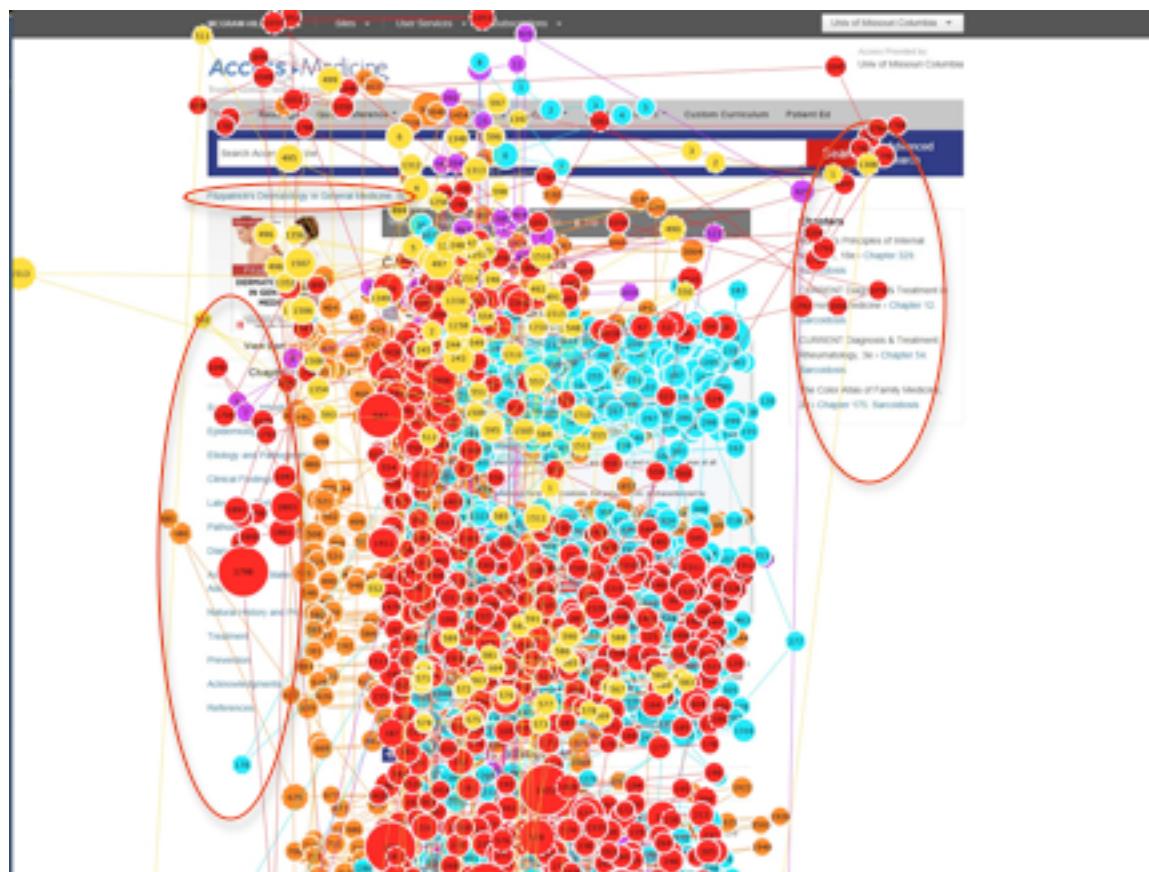
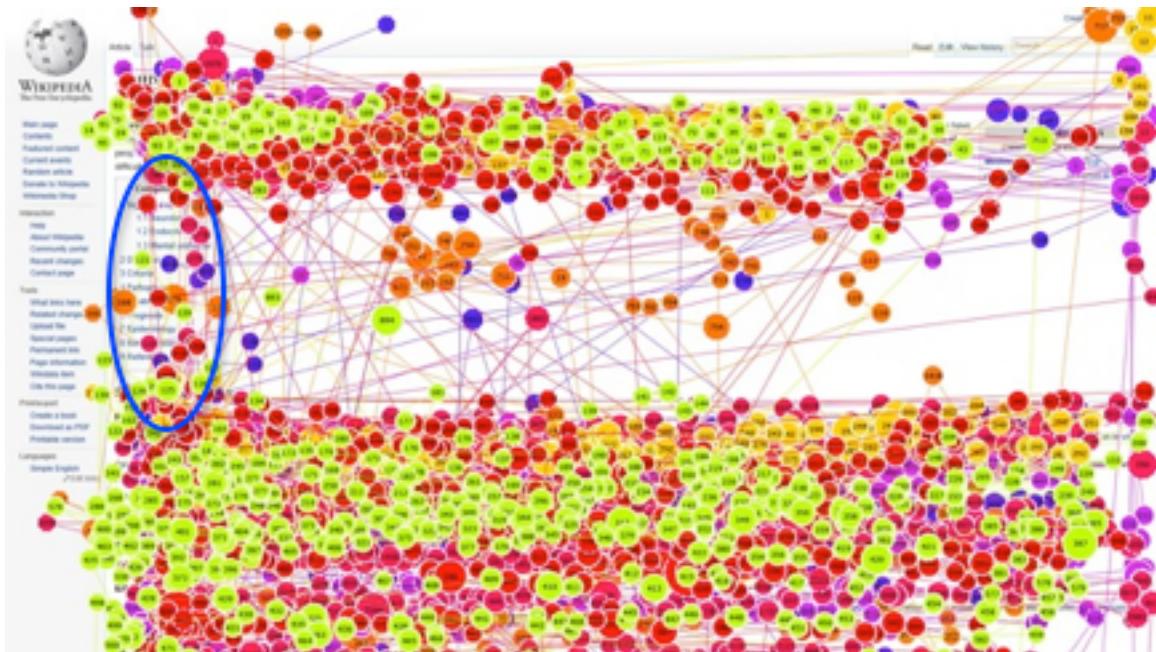


Figure 29. Example of student interactions with navigation in *Access Medicine*

Analysis of participants' eye metrics showed that, on average, they looked at the outline located on the left about three times ( $M = 3.16$ ,  $SD = 5.32$ ) for a total of 1.24 seconds ( $M = 1.24$ ,  $SD = 2.03$ ) but almost never clicked on it ( $M = 0.5$ ,  $SD = 1.42$ ). Students also almost never looked at or clicked on the related content (total visit duration ( $M = 0.45$ ,  $SD = 0.85$ ), visit count ( $M = 0.83$ ,  $SD = 1.42$ ), and mouse click count ( $M = 0$ ,  $SD = 0$ )). None of the participants clicked on the breadcrumb navigation trail for navigation purposes (mouse click count = 0) (Figure 29).

In *Wikipedia*, the table of contents served as the navigation menu (Figure 30). Participants on average looked at the table of contents about six times ( $M = 5.89$ ,  $SD = 3.5$ ) for a total of about 2 seconds ( $M = 1.98$ ,  $SD = 1.75$ ); however, almost never clicked on it ( $M = 0.21$ ,  $SD = 0.53$ ). Such infrequent interactions with the navigation menu could have been the result of its placement on the page – once the users scrolled down the page, navigation menu remained above the page fold and, hence, was invisible and unavailable.



*Figure 30.* Example of student interactions with navigation in *Wikipedia*

When asked about the utility of navigation elements in all three information resources, participants commented that the navigation menu in *UpToDate* was the most useful. It allowed them to quickly jump to the needed section: “*There is stuff that I don’t necessarily use, but it’s laid out in such a way that I can easily skip over that section if I want. Outline that is on the left side lets you click on the each one [hyperlink], so if I wanted to look over treatment for some reason and I’m trying to figure out the pathology, I can just go straight to that.*” Students also admitted that instead of using the navigation

menu in *UpToDate*, they often used *CTRL + F* key combination to find a specific piece of information.

### **Students' Perceptions of the Appropriateness of Information Resources for Learning**

**UpToDate.** Students named *UpToDate* as one of the ‘go to’ resources they used for learning because it contained the most accurate, up-to-date, and in-depth medical information. The majority of participants liked using *UpToDate* in everyday learning because it assimilated and synthesized information from a lot of reputable resources in one place. One participant commented: “*I really like it [UpToDate] because it brings together the primary sources. It would take me forever to go through every article and read the data. It kind of does it for you. Whereas with a textbook, you really don't get that.*”

Ten students commented that *UpToDate* contained “*literally every single thing they could think of a disease*”, which sometimes was too much for pre-clerkship medical students’ learning needs. For instance, for their clinical exams and PBL objectives students needed to know one or two of the most efficient and widely used treatments. When they used *UpToDate* to find this information, they were presented with all treatment options available for that medical condition, including clinical trials that were not used as a definitive treatment. Abundance of options to choose from made it more difficult for students to make their choices and made them feel less confident in their answers. As one student commented: “*When I'm looking for information, I need to know which drug is used for treatment and if there is good evidence, that is good enough for me. I don't really need to know where the trial was conducted. Like in Netherlands.*

*UpToDate goes into the details too much. And I see why they do that because that's what the clinicians use it for. But at my level, we are never tested on that."*

Students explained that background information was not available in *UpToDate* as its users were expected to have background knowledge of the topic. One student commented: "*UpToDate always says you should do a biopsy, but they don't tell me what I should look for in the biopsy. So I need to go read an additional article on UpToDate about biopsy, which is kind of annoying*". As a result, to approach the richness of information on *UpToDate* with confidence, students first needed to obtain sufficient background knowledge of the topic.

Fifteen participants were in agreement that due to lack of background information in *UpToDate*, this resource was better suited for learning about clinical aspects of medicine, e.g., diseases, drugs, diagnostics, treatment, and management: "*It [UpToDate] uses the latest research that is presented to clinicians, so that they can perform the best practice and is not really designed for medical students in mind at all.*" Still, several students were of the opinion that they would rather see more information than less of it.

Clinical orientation of *UpToDate* also resulted in a lot of medical terminology unfamiliar to first-year medical students. Some students commented that during the experiment it was difficult for them to read only from *UpToDate* and understand the processes of diseases because they were not familiar with the terminology and concepts and had no way of learning about them in more detail first: "*It [UpToDate] was catered to more practicing physicians and all the terminology used in there, so I had to do a lot of thinking with that one.*"

Other students noted that it was difficult for them to read from *UpToDate* not only because of unfamiliar terminology but also because of the layout, e.g., the main text was too dense and stretched on the page, the headings were not big enough, and there was not enough partitioning of information into sections. One participant commented: “*... for neurosarcoidosis UpToDate was very dense. You kind of reread the paragraph and you are like, I didn't really comprehend any of that, so you go back and read it again a couple of times. I found myself doing this a couple of times for neurosarcoidosis.... So if you're not really dialed in and focused on what you're reading, you can get lost really, really quickly.*”

Some students noticed grammatical complexity of the sentences that interfered with the efficiency of information acquisition: “*I think it is sentence structure, putting more than one complex idea in one sentence. On Wikipedia page it has one idea - one sentence, next idea - next sentence. It [Wikipedia] is more simplified as far as the structure. Whereas some other medical texts [on UpToDate], they compound different ideas into one sentence and you have to really be paying attention or read multiple times to get all those ideas.*”

Participants described a few strategies as a way of dealing with difficulties of information acquisition from *UpToDate* due to layout constraints. For example, in order not to lose the line during reading, some students highlighted sections of the text, which helped them focus. Others used the mouse cursor to tag along the text as they read it. The third re-sized the browser window to make it long and narrow, which allowed them to read from top to bottom rather than across the page. The fourth zoomed in on the text to a maximum. The fifth relied a lot on the left-side navigation to navigate through the

content. The sixth extensively used the summary of the article provided at the end of the page to gain an idea of the content of the article and decide on its appropriateness (Figure 31). Others read summary of the topic located at the end of the page to solidify understanding of the article: “*UpToDate is really, really in depth. That sometimes, kind of, slows me down as well because you can get lost in emerging treatments and everything. Sometimes I will read that article and at the end, it always has the summary. I always read it to make sure that I remember reading what they are talking about and that it makes sense*”.



Figure 31. Example of reading the summary in *UpToDate*

**Access Medicine.** Among the three information resources, *Access Medicine* was named the most appropriate for learning in pre-clerkship curriculum. Twelve students

were ok with the fact that *Access Medicine* did not contain the most recent information – they believed that scientific concepts did not need to be updated too often.

Students were of the opinion that *Access Medicine* was written easier than *UpToDate*, as if “*someone was talking to them*”, with easier grammatical structures that facilitated comprehension. Some students suggested that the ease of *Access Medicine* use could have been due to its organization as a textbook, which reminded them of paper textbooks used during their undergraduate studies. On the other hand, *Access Medicine* was reported to have a large number of acronyms, e.g., UIP, PVT, FAB, which impeded understanding of the content. The acronyms were usually explained once at the beginning of the chapter and were not hyperlinked. That is why, when students encountered acronyms further down in the text, the only way to look them up was to scroll all the way up and find their initial explanations.

Overall, students liked that they could have access to multiple textbooks gathered in one place when using *Access Medicine*. Ability to look around and use multiple resources made some students feel like they could get a very comprehensive overview and understanding of the topic: “*The nice thing about Access Medicine is different books take different spins on things, and the obstetrics book on the flu is going to be different from emergency medicine on the flu. So it's nice to use it for different cases. That's the nice part about Access Medicine, you can get information on different points of view within medicine.*”

Students’ opinions about information organization in *Access Medicine*, however, varied. Some students liked the fact that information was organized by textbook because they knew which book or chapter they were using at the moment. Others found such

information organization disjointed and time-consuming: “*I felt like I was jumping around a lot and could not just sit and read through all of it in one spot.*” These students felt like because search results were organized by the book, the decisions about the relevancy of the search had to be based on the book rather than the topic, which made the search experience less efficient.

Students also could not agree on the usefulness of drop-down/expandable sections that each chapter in *Access Medicine* was divided into (Figure 32). Some students really liked being “*restricted*” within one specific section; this made them feel like they knew exactly what they were reading about and how much they had to read. Others felt frustrated that they needed to expand each section separately in order to gain access to its content: “*Having to click on each section separately for information to expand is kind of irritating. I'd rather it just be there and I can click on one thing rather than expand and change – just have it all pop up at once... I would rather scroll down than have to go click on each section and wait for it to pop up.*”

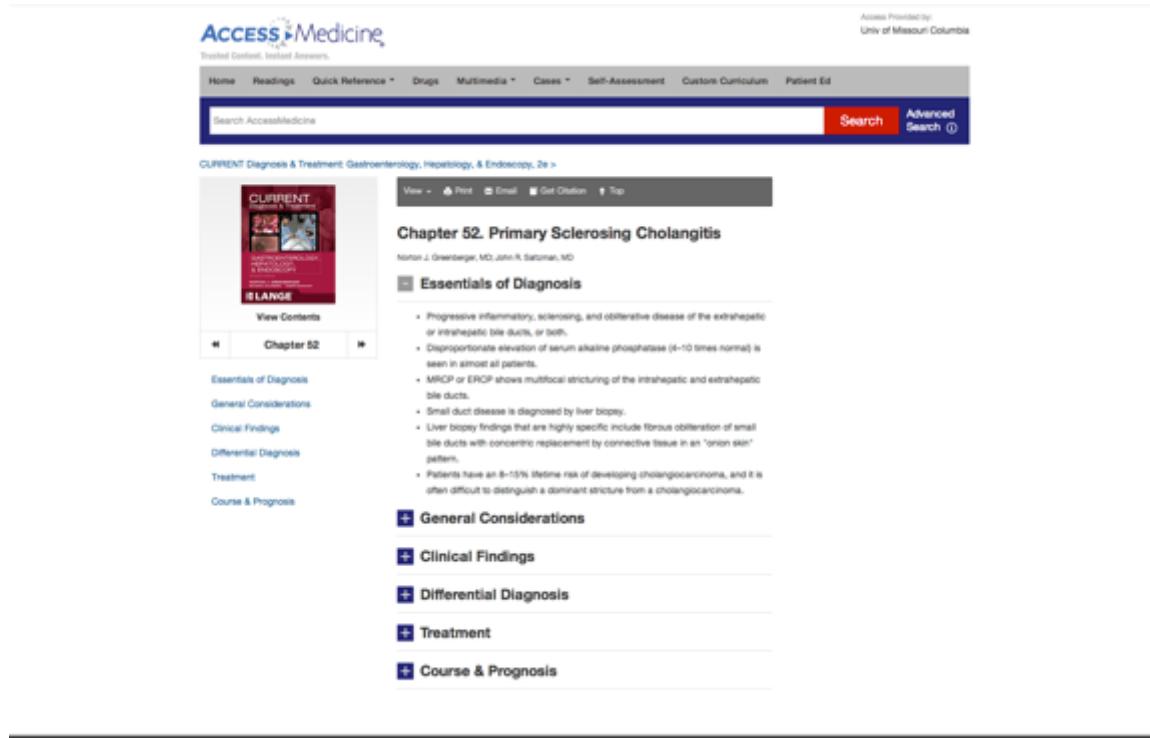


Figure 32. Drop-down sections in *Access Medicine*

Another point of disagreement was the presence of wide white space on both sides of the main content. Some students believed that such outline made the interface look clean, less cluttered, and easier to read. Others did not like the text being in the center with too much white space around it because such layout made the text look small on a large screen, even when students zoomed in on it. During the interviews participants added that the current layout of *Access Medicine* was the result of its recent interface redesign. Previously, *Access Medicine* had no white space around the text and search by the topic was by default. A few students admitted that since the change, they completely stopped using *Access Medicine* because discomfort they experienced with information access hampered their learning.

**Wikipedia.** During the interviews students commented that *Wikipedia* played an important role in their learning. Except for the two participants, the rest unanimously

confirmed they always began learning of a new topic in *Wikipedia*. In this role *Wikipedia* was found extremely helpful for obtaining a background understanding of a topic in a quick and easy way and determining the direction of further readings and research. As one student commented: “*Wikipedia helps you to narrow down and focus what you need to be reading about. It’s also like a big picture in your head and you can go and fill in the details once you get to the other sources*”.

Students particularly liked the following *Wikipedia* features for the following reasons:

- 1) Effective information organization and presentation such as readable font, visible headings, clearly demarcated sections, and plenty of hyperlinks contributed to quick and easy information acquisition and processing during reading.
- 2) Topic summary at the beginning of every article helped students obtain a general idea of the topic and decide whether or not the article was relevant.
- 3) Images that were available on the same page as content sometimes served as additional sources of information.

Several students warned that simplicity of information presentation on *Wikipedia* could have contributed to a “*false sense of security*”, i.e., students were confident in their knowledge and understanding of diseases as they were reading from *Wikipedia*; however, during the test, they realized they did not acquire enough new knowledge and could not meet the difficulty of the test questions. Others noted that simplicity of information presentation made them feel like they knew a lot about the topic, and filtering out familiar content slowed down their information acquisition.

When using *Wikipedia* for learning, participants found it to be very well written, accurate, and up-to-date. Four students believed they did better on the test after using *Wikipedia* than after using *UpToDate* or *Access Medicine*. Several students commented that, based on their experience, *Wikipedia* content was always in line with what they read in other information resources and they never had any bad luck with *Wikipedia*. At the same time, students were aware that *Wikipedia* content could be generated by anyone; that is why, they stayed alert when using this information resource. In some instances, students established trustworthiness of *Wikipedia* content by checking the list of references at the bottom of the article. If those references came from reputable sources and could be retrieved, students tended to perceive the content of *Wikipedia* article as trustworthy.

## CHAPTER 5 – DISCUSSION

### Research Question 1

What are the differences in learners' mental effort invested in interactions with electronic information resources?

In order to answer the first research question, I investigated how learners' mental effort differed when they interacted with the three electronic information resources of interest, namely, *UpToDate*, *Access Medicine*, and *Wikipedia*. Mental effort was measured through the eye movement metrics such as total fixation duration, fixation count, total visit duration, and task-evoked pupillary response.

The cognitive load theory assumes that intrinsic and extraneous cognitive loads are additive (van Merriënboer & Sweller, 2010). This means that when intrinsic cognitive load is high, extraneous cognitive load needs to be low to ensure efficient learning and vice versa. For this study, intrinsic cognitive load of the learning material was intentionally maintained high. This provided an opportunity to focus on the sources of extraneous cognitive load (e.g., material presentation and interface design characteristics of electronic information resources) as contributing to overall cognitive load.

The results of the repeated Latin Square Design analysis indicated that learners' mental effort, when measured through total fixation duration, fixation count, and total visit duration, was significantly different when they interacted with the three electronic information resources. Specifically, mental effort values were significantly higher when learners interacted with *UpToDate*. Below are a few explanations of this finding.

Geissler, Zinkhan, & Watson (2001) and Miniukovich & De Angeli (2014) proposed to consider interface design elements of websites such as large text blocks,

lengthy pages, density of information, and the number of hyperlinks as contributing to subjective user perceptions of visual complexity. The findings of the current study suggested that students perceived *UpToDate* interface as overwhelming and complex due to abundance, density, and complexity of available information. Due to such way of information presentation, students often had to scroll down the pages in search of the needed content and/or heavily rely on the navigation menu for orientation on the page. It has been shown that user interactions with content that is not optimized to fit the monitor window tend to increase their mental workload (Wästlund, Norlander, & Archer, 2008). It is, therefore, possible that in addition to overcoming visual complexity of *UpToDate* interface, students' mental effort was also expended on locating the needed content.

It is worth mentioning that in hypertext, in addition to material presentation, extraneous cognitive load can also be imposed by the reading processes that include letter recognition and decoding words (Antonenko & Niederhauser, 2010). During the interviews several students commented that in order to find needed information in *UpToDate*, they had to browse through a large amount of text, which often made the process of information acquisition unnecessarily time-consuming. It is likely that interacting with more text resulted in more reading and eye movement activity, which, in its turn, was reflected in higher values of total fixation duration, fixation count, and total visit duration when using *UpToDate*.

Access to information in *UpToDate* was further complicated by semantic and structural complexity of the sentences. Paterson, Liversedge, & Underwood (1999) and Rayner, Carlson, & Frazier (1983) showed that when readers initially misanalyse syntactically ambiguous sentences, they tend to spend more time on the disambiguating

word. It is possible that spending more time trying to understand complex sentences and unfamiliar words resulted in the increase in fixation duration and fixation number in situations when participants had to re-read semantically complex sentences in *UpToDate*.

Design of hyperlinks in *UpToDate* deserves a special attention in explaining the peculiarities of students' information acquisition behavior in this resource. There were very few one-word hyperlinks but a lot more sentence-long hyperlinks. In quite a few instances, students perceived these lengthy hyperlinks as obstacles to the continuity of reading and efficient information acquisition. These links, although informative, were almost never followed, because students did not want to be lost in the dense content of the articles and spend additional time reinstating the ideas in their memory. It is, therefore, possible that trying to intentionally avoid interaction with this type of hyperlinks during the process of meaning construction required unnecessary mental effort from learners and interfered with text comprehension. The Cognitive Theory of Multimedia Learning refers to interesting but irrelevant information as extraneous cognitive load (Mayer, 2005) and recommends 'weeding it out' for the purpose of supplying the learner with essential information in a short and concise manner (Mayer & Moreno, 2003).

When talking about cognitive overload, it might be important to know which of the three types of cognitive load is mostly contributing to it. In the analysis of their studies Gerjets & Scheiter (2003) showed that depending on the goal configurations (e.g., teacher goal vs learner goals) and information processing strategies (e.g., comparing and contrasting the information vs elaborating on the found information), learners demonstrated fluctuations in germane and extraneous cognitive load during the tasks. The

authors explained their findings with the help of the theories of action control, which postulate that there is a close relationship between task difficulty and the amount of effort invested in the task (Kuhl, 1984). Specifically, the difficulty of the task leads to an increased effort on the part of the learner (e.g., attention, concentration), which, in turn, helps overlook the sources of unnecessary extraneous cognitive load.

Above I elaborated that I believe that in the case of *UpToDate* the increase in mental effort occurred primarily due to increase in extraneous load imposed by the format of material presentation and interface design characteristics of this information resource. It is, however, possible that increase in mental effort when using *UpToDate* occurred due to increase in germane cognitive load. In accordance with the principles of the action control theories (Kuhl, 1984), it is possible to assume that when students interacted with *UpToDate*, their mental effort was directed towards understanding the difficult material instead of overcoming the difficulties associated with its presentation. As a result, students' short memory was preoccupied with germane rather than extraneous cognitive load, which was reflected in their eye movement characteristics. Since differentiating between the types of cognitive load was not part of this study, it remains unclear whether increase in mental effort was due to the increase in extraneous or germane cognitive load.

Another measure of mental effort implemented in this study was task-evoked pupillary response (TEPR). It included the measurements of learners' mean pupil dilation, lability, and percentage change in pupil size when interacting with information resources. Results of the repeated LSD showed no statistically significant difference in TEPR values among the three information resources. What was interesting, though, was

the fact that students' interactions with all three information resources were characterized by pupil constriction rather than dilation from baseline values.

Although it is most common to observe pupil dilation in response to increased work load (Iqbal et al., 2004), it has been suggested that pupil constriction can also be an indication of high information processing demands (Poock, 1973). Based on this suggestion, I believe that students' interactions with the three information resources during the experiment were characterized by high intensity of information processing close to their maximum cognitive capacity. Conditions of the experiments could have contributed to the intensity of information processing since students had a limited amount of time (~20-25 minutes), during which they needed to learn a lot about each disease. In the end, the fact that participants' pupil constricted from baseline by about 2% more when they acquired information from *UpToDate* than the other two information resources suggested that learning from *UpToDate* was associated with higher cognitive demands and mental effort.

Elaborations on the applied eye-tracking methodology once again confirmed previous findings that this approach to data collection results in very rich data that requires time-consuming and labor-intensive analysis (Hyönä, 2010; van Gog & Scheiter, 2010). Students' feedback about *UpToDate* being the most difficult information resource to use and acquire information from complemented the identified eye movement patterns, confirming the highest mental effort invested in interactions with this information resource. This finding suggested that application of eye tracking in this study provided accurate and informative measures of cognitive processes.

## **Research Question 2**

How do the differences in learners' mental effort invested in interactions with electronic information resources affect learning?

In order to answer the second research question, I investigated whether the differences in mental effort affected learning and if so, in what way. In this study, learning was evaluated with the help of the three ten multiple-choice-question knowledge tests. Each test was conceived to accurately measure comprehension of each of three diseases and was offered immediately after students finished reading about each disease.

Sweller (1994) argued that a combination of high intrinsic and extraneous cognitive load may overwhelm the limited capacity of the working memory needed for constructing meaning and thus impair learning. This study's findings suggested that *UpToDate* was associated with a significantly higher amount of mental effort, likely to have been caused by high extraneous cognitive load. According to the assumptions of the cognitive load theory, it was expected that high intrinsic cognitive load of the learning material and high extraneous cognitive load associated with the use of *UpToDate* would result in inferior learning from this information resource. The repeated Latin square design analysis, however, yielded no statistically significant difference between learning outcomes and the amount of mental effort invested in information resources. In other words, even though *UpToDate* was associated with significantly higher mental effort, learning outcomes after using *UpToDate* were not any less inferior or more superior than after using *Access Medicine* or *Wikipedia*. Furthermore, findings from the descriptive statistical analysis of the test results showed that on average participants were able to answer five out of ten questions correctly regardless of the information resource used.

This particular finding suggested that despite the differences in information acquisition patterns and the amount of acquired information, students were on average consistent in the amount of knowledge they gained from each of the three information resources. This finding also served as a validation of the appropriateness of the knowledge test and the three diseases chosen for the experiment. In other words, the fact that students were able to answer correctly on average fifty percent of the test questions regardless of the disease and the type of information resource used suggested that the difficulty of the diseases and, hence, intrinsic cognitive load associated with learning, was pretty much on the same level.

Results indicated that differences in mental load associated with differences in design of the information did not make a difference in learning. However, difference in learning could have been due to student individual experiences, like satisfaction and frustration. In addition, lack of statistical difference in learning outcomes could have been due to the fact that participants were believed to have a more or less similar level of prior knowledge. Finally, had the participants had less time to complete the tasks, there may have been differences related to learning outcomes, possibly as a result of different designs.

In the absence of statistically significant difference between the amount of mental effort invested in information resources and learning, I proceeded to explore this question by looking for factors that could have contributed to the achievement of similar performance. Specifically, I referred to observations of students' interactions with information resources (particularly, the role of hyperlinks) and their feedback obtained during the interviews, which are discussed in detail in the next section.

### **Research Question 3**

How do certain interface design characteristics of electronic information resources contribute to or interfere with information acquisition and, consequently, learning?

In this section I discuss findings with regard to a few major interface design characteristics of the studied electronic information resources. I also discuss how these features could have affected the processes of information acquisition and learning.

#### **The Search Page**

Only *UpToDate* and *Access Medicine* had separate search pages that listed all returned results based on the performed search. There was, however, a big difference in the design of these pages, which possibly was the reason why students spent a different amount of time on these pages. Specifically, students spent almost twice as long on the search page in *Access Medicine* than in *UpToDate* possibly for the following reasons. First, in *Access Medicine* search results were organized by the book; that is why, when students searched for a disease, they had access to book chapters that contained excerpts of information on that disease instead of comprehensive description of the disease. Such organization of the search often resulted in the same book (although different chapters) appearing several times on the search page (i.e., redundancy of information).

Second, organization of the search results by the book did not yield any additional details with regards to the topic, i.e., students had no way of knowing whether the book chapter contained clinical manifestations, diagnosis, or differential diagnosis of the diseases. It is possible that the lack of additional, more detailed information about the searched topic slowed down and complicated students' decisions with regards to textbook selection. Not being able to make more informative decisions, students often scrolled up

and down the page in search of the books with familiar cover, i.e., the books they used before.

Third, observations of student searching behavior demonstrated that they hardly ever used filters available to them on the search page intended to help narrow down the search by topic, the type of reading (e.g., textbooks, diagnostic tests, guidelines) or the type of multimedia (e.g., interactive modules, videos). During the interviews students commented that *Access Medicine* recently underwent a re-design. It is possible that students were not yet accustomed to the new features and their placement.

## **Navigation**

Analysis of students' behavior showed that during the tasks participants' interactions with the additional navigation elements in the three electronic resources were infrequent. This was an interesting finding considering that design and placement of navigation elements in *UpToDate*, *Access Medicine*, and *Wikipedia* were very different. It is possible that low use of navigation elements was the result of the nature of the assigned task and time limitations, which made students focus more on reading the content rather than exploring different options. However, it is also possible that it was the placement of navigation elements on the pages that prevented students from their use. For example, it was interesting to observe that students did look at the left-side navigation menu in *UpToDate* quite a lot but hardly ever clicked on it. Perhaps, they chose to orient themselves within the text by scrolling up and down the page rather than using hyperlinks available in the navigation menu.

In *Access Medicine*, related content located on the right side on the page was only visible above the page fold. The book outline located on the left side of the page was not

affected by the scroll or page fold; however, it only provided access to other book chapters. If students were interested in learning about different aspects of a disease, titles of book chapters did not allow them to know whether these books contained needed references. Similarly, in *Wikipedia* table of contents was only visible above the page fold; when students scrolled down the page, article content was not visible to them any more and could not be interacted with.

### **Character/background color combination**

Analysis of interfaces indicated that the color of hyperlinks differed among the three resources. In *UpToDate*, hyperlinks appeared in dark green color on the white background in between the black color of the main text. In *Access Medicine* and *Wikipedia*, hyperlinks were dark and bright blue accordingly on the white background in between the black color of the main text.

As found by Bhattacharyya, Chowdhury, Chatterjee, Pal, & Majumdar (2014), blue as character on a light background was associated with lower cognitive load, lesser searching time, minimum error, and optimum eye movement activities. It is possible that the blue color of hyperlinks in *Wikipedia* along with their extensive quantities made them stand out against the rest of the text and, therefore, be more noticeable in students' peripheral vision. By being able to clearly distinguish between the main text and hyperlinks, students likely could tell in advance which hyperlinks they were going to follow and which once would be helpful in planning their navigation.

In *UpToDate* and *Access Medicine*, the color contrast between hyperlinks and main text was not very prominent. Interestingly, Bhattacharyya, Chowdhury, Chatterjee, Pal, & Majumdar (2014) also found the error rate during onscreen searching tasks was

the second highest with green character/white background combination. It is possible that darker color of the hyperlinks next to the black color of the main text did not make hyperlinks in these information resources stand out. As a result, students possibly did not get the chance to explore some of the hyperlinks that could have been helpful in their learning.

### **Conformance to the principles of multimedia learning**

In this study, the chosen learning material was considered to be of high intrinsic cognitive load as the participants had low or no prior knowledge of the diseases. In such circumstances, several interface design characteristics of electronic resources were found helpful in manipulating the amounts of high intrinsic cognitive load. According to the segmenting principle of multimedia learning, people learn more effectively when they have the ability to interact with material presentation and control its pace (Mayer et al., 2003). This principle was very clearly supported in the design of *Access Medicine* that allowed learners to expand chapter sections one at a time. Some students found such content presentation useful because they could focus on the limited amount of information available at one time and transition between the sections at a speed convenient to them, thus, regulating the amount of experienced intrinsic cognitive load and making it more manageable. Others, however, did not find this feature particularly useful.

According to another principle of multimedia learning – signaling – people learn better when they are presented with cues that emphasize the organization of the material and help activate learners' knowledge structures (Mayer & Moreno, 2003). This strategy is proven especially helpful when learners try to select relevant information in order to

integrate it with prior knowledge. An example of an organizational cue is a heading. It has been argued that signals or cues, like topic heading, have greater benefits for text recall, especially when the structure of the topic is relatively complex (Lorch, Lorch, & Inman, 1993) and the learners do not have prior knowledge of the topic (Lorch & Lorch, 1996). Furthermore, Hyönä & Lorch (2004) found that headings are among the first text elements that readers notice when looking for relevant information, which is reflected in their eye movement patterns.

From the three evaluated information resources only *Access Medicine* and *Wikipedia* were found to have headings that clearly stood out against the rest of the content and were reported helpful during hypertext navigation. Headings in *UpToDate* were not as noticeable or helpful for understanding the organization of the content.

The fundamental principle of multimedia learning states that people learn from words and pictures better than from words alone (Mayer, 2005). In addition, people learn better when corresponding words and pictures are presented near rather than far from each other (e.g., spatial contiguity principle) or simultaneously rather than successively (e.g., temporal contiguity principle) (Mayer & Moreno, 2003). Further, through the analysis of eye movement, Holsanova, Holmberg, & Holmqvist (2008) showed that close proximity of graphics and related text makes it easier for users to integrate information. This principle was not supported in the case of *Wikipedia* and *UpToDate*. Specifically, analysis of their interfaces showed that images in *Wikipedia*, although present on the pages with the text, were not referenced in the text. This made *Wikipedia* images somewhat disjointed from the text. In *UpToDate*, images were not present on the pages with the textual information at all. The images appeared in a separate window after a

hyperlink leading to them was activated. It is possible that such placement of images in relation to the text was in violation of spatial and temporal contiguity principles causing split-attention effect and not contributing to learning. In addition, because students had to keep the semantic connections between the text and the image in their memory, they relied more on recall rather than recognition of information, which was in violation of one of the usability heuristic principles (Nielsen, 1995). .

In an attempt to improve learning in hypermedia, Hoffman (1997) proposed to rely on the principles of Elaboration Theory (cf., Reigeluth & Stein, 1983) for content organization and presentation. Specifically, she proposed that in accordance with this theory, at the very beginning of the page, hypermedia learning environments should include overviews presenting learners with a few general but fundamental ideas about the topic. In situations of high intrinsic cognitive load, these overviews could guide learners in their decisions as to which information to focus on in their further readings.

Review of the interfaces indicated that both *Wikipedia* and *UpToDate* contained examples of the proposed overviews. The difference between them, however, was in their placement. In *Wikipedia* the overviews were located at the beginning of the page, while in *UpToDate* overviews were placed at the very bottom of the page. It is possible that the differences in placement of overviews on the page could have contributed to the differences in students' navigational decisions and the efficiency of information acquisition from these two information resources.

In their investigation of leads (or brief summaries or overviews of information) Antonenko & Niederhauser (2010) found that the use of leads also helped regulate cognitive load. Specifically, by helping orient the reader to the linked information, leads

helped increase germane cognitive load and minimize mental load associated with establishing the connection between the nodes. At the same time, leads helped reduce extraneous cognitive load by aiding to overcome the split attention effect caused by reading content of two or more different nodes. It is possible that learning with *Wikipedia* was enhanced to a certain degree by students' extensive use of overviews.

## Hyperlinks

**Hyperlinks in *UpToDate*.** One of the great advantages of hypertext learning is freedom of navigation that supports the exploration of available information. Analysis of the *UpToDate* interface revealed that the number of hyperlinks throughout the main content was rather scarce. Some of these hyperlinks were a one- or two-word hyperlinks; others were lengthy, sentence-long hyperlinks. Hyperlink functionality in *UpToDate* was reported to be inconsistent and unpredictable. Namely, participants often commented that they were not aware of where certain hyperlinks took them, i.e., either to a different spot within the same article or to a new article.

According to the participants' feedback and observations of their interactions with hyperlinks in *UpToDate*, design and functionality of hyperlinks could have prevented students from their efficient use and, possibly, negatively affected learning. In the conducted experiment, information acquisition was limited to 20-25 minutes. In such circumstances not knowing where exactly students would land had they decided to follow certain hyperlinks and/or having no opportunity to quickly look up unknown words with the help of hyperlinks resulted in passive or no use of embedded hyperlinks. It is also possible that by minimizing hyperlink use in *UpToDate*, students tried to avoid situations of being lost-in-hyperspace and wasting the task time on finding their ways back. As a

result, instead of exploring additional related information, students opted to stay within the main article and focus on extracting meaning from it.

Analysis of eye metrics during interactions with *UpToDate* supported the hypothesis of students' increased attention and concentration. Findings from the research question one showed that *UpToDate* was associated with significantly longer fixation duration and significantly higher number of fixations. Togami (1984) found that these eye metrics were indicative of higher attention and accuracy of visual inspection and, as a result, lead to improved performance. This, however, was not the case with *UpToDate*. Despite the fact that students paid more attention to the content, their learning outcomes after using *UpToDate* were not different from those achieved after using the other two information resources.

**Hyperlinks in Wikipedia.** Information acquisition from *Wikipedia* was associated with significantly lower mental effort compared to *UpToDate* and resulted in the same quality of learning as from the other two information resources. I choose to believe that *Wikipedia* interface design characteristics were responsible for such outcome. First, students commented that information in *Wikipedia* was accessible to them and easy to comprehend. Second, navigation was very straightforward: there was no additional effort invested in decisions as to which hyperlinks to follow and where they would take the students. Students followed hyperlinks based on their immediately arising information needs to quickly look something up, address their knowledge gap, and assist with more complex information searching if necessary – the finding similar to the one reported by Lim (2009).

When describing hypermedia environments, Conklin (1987) warned against the

dangers of experiencing cognitive overhead caused by decisions as to what hyperlinks to follow. DeStephano & LeFevre (2007) added that these decisions could be complicated even more by the greater number of hyperlinks and their type, ultimately negatively affecting the quality of learning. Analysis of *Wikipedia* interfaces revealed that it had visibly more hyperlinks embedded in each page. Observations indicated that students took advantage of hyperlinks in *Wikipedia* and extensively used them throughout the task, reporting no additional cognitive effort invested in their selection and use based on their quantity.

More extensive use of hyperlinks in *Wikipedia* resulted in a greater number of visited individual pages. Throughout the three tasks, each participant visited an average of 15 pages in *Wikipedia* compared to 12 pages in *Access Medicine* and 10 pages in *UpToDate*. It is likely that by visiting more pages and timely addressing their information needs when using *Wikipedia* students were able to engage in deeper cognitive processes by constantly comparing and contrasting newly found information with existing knowledge. In the context of hypertext learning, Spiro, Coulson, Feltovich, & Anderson (1988) called such behavior ‘criss-crossing the conceptual landscape’ and Jacobson & Spiro (1995) suggested that it promotes learners’ cognitive flexibility. It is also possible that active and purposeful use of hyperlinks in *Wikipedia* promoted active knowledge construction (Kintsch, 1998) and allowed learners to establish connections between the concepts and their further integration into existing knowledge.

When talking about non-linear information access and the need to plan navigation inherent to hypermedia environments, Zumbach & Mohraz (2008) and Zumbach (2006) referred to these types of decisions as cognitive overhead. They argued that cognitive

overhead could actually lead to deeper, more conscious information processing and enhance learning, because learners get exposed to a greater variety of information. More so, cognitive overhead can be especially advantageous for difficult tasks as it allows to meet the complexity of ill-structured domains such as medicine (Zumbach & Mohraz, 2008). It is possible that in the case of *Wikipedia*, cognitive overhead caused by a large number of hyperlinks was, in fact, beneficial for learning. Extensive use of hyperlinks contributed to acquiring a broader picture of the diseases and allowed learners to compensate for the potential lack of depth and quality of information available in *Wikipedia*, which was reported previously (e.g., Azer, 2014; Leithner et al., 2010). Similar to Hilligoss & Rieh (2008), I believe that accessibility of the content and abundance of navigational options made *Wikipedia* a suitable resource for obtaining new ideas and perspectives on a new topic.

In their study Paas, Van Gerven, & Wouters (2007) described an experiment, during which one of the compared groups achieved the same test performance but with less mental effort due to different levels of instructional efficiency. Similarly, I conclude that participants of this study achieved similar learning outcomes but with significant differences in the invested mental effort: less mental effort was invested in interactions with *Access Medicine* and *Wikipedia* and more mental effort was invested in interactions with *UpToDate*.

Overall, reported ease of comprehension and use of information available in *Wikipedia* in combination with the achieved learning outcomes made me agree with previous research (e.g., Fallis, 2008; Wood & Struthers, 2010; Haigh, 2011) in concluding that in this particular case *Wikipedia* was no less inferior information resource

for learning than reputable resources like *UpToDate* and *Access Medicine*. This is a very noteworthy finding because in this particular case it suggests that learning after using *UpToDate*, *Access Medicine*, and *Wikipedia* were of similar quality. This finding makes a significant contribution to the ongoing debate as to the appropriateness of questionable information resources like *Wikipedia* for learning.

**Hyperlinks in *Access Medicine*.** Due to very few hyperlinks in *Access Medicine*, their use during the task was minimum and unremarkable. It is possible, though, that the scarcity of hyperlinks in *Access Medicine* was counterbalanced by a wide array of textbooks that each provided a unique perspective on the disease. At the same time, shortage of hyperlinks in *Access Medicine* did not allow students to smoothly transition between the textbooks and related content. Students had to return to the main search page in order to select a new textbook even though the disease remained the same. Inability to connect information on the same topic from several textbooks at once possibly made students spend additional time on navigating back and forth to the main search page instead of learning the material (~ 45 seconds total throughout the task). Such interruptions in the information acquisition were possibly the reason why the number of “I don’t know/I don’t remember reading about it” answers was the highest in *Access Medicine*.

### **Limitations, Implications, and Future Research**

#### **Limitations**

Despite noteworthy findings, the results of the study should be interpreted with caution due to a number of limitations. First, during the tasks students were limited in their interactions with just one information resource at a time. This was not a natural

occurrence for them as students reported usually using a combination of information resources for learning. In addition, even though all study participants reported good familiarity with the three information resources, they also commented on the differences in their preferences for these resources. Both factors could have affected students' use patterns of information resources.

Second, even though the three diseases that served as a learning material were carefully selected and approved for the study, some may believe that the element of subjectivity in their selection was present. Despite the fact that the panel of experts found all three diseases to be of similar complexity, there is a chance that students perceived these diseases differently, for example, because diseases were pertinent to different organs and body systems.

Third, during the tasks of information gathering, people usually require prolonged task completion time, during which they may study certain things multiple times (Kellar, Watters, & Shepherd, 2007). In this study students were given a limited time to learn about each disease (~ 20-25 minutes), which was not their natural behavior. Still, for the purpose of the experiment I considered such duration of the tasks to be realistic and acceptable, especially since I only needed a sample of students' information acquisition behavior.

Fourth, to observe the validity of the eye tracking data, it was impossible to zoom in/out on the content available in information resources. As a result, students had to work with the font that was smaller than they were used to, which could have added additional strain on their eyes during reading.

Fifth, even though the unit of analysis included 54 student interactions with information resources, the sample size of 18 participants could have been considered relatively small. Increasing the sample size could have resulted in increased variability of the knowledge test results.

In view of several limitations, it is cautioned that all findings should be considered as pertinent to this study only. It is also advisable not to generalize the findings beyond the context of this study.

### **Implications**

Based on the findings, the study carries implications for a number of user groups. First, it is suggested that medical school faculty and students carefully select certain information resources for learning. Specifically, *UpToDate* may be too cognitively demanding for first-year medical students when learning about a new topic; *Wikipedia*, even though easy to use and understand, may be less accurate; and *Access Medicine*, even though primarily intended for learning, may be outdated. Therefore, it is recommended to use a combination of these resources to effectively address information and learning needs.

Second, instructional designers need to take into account the principles of cognitive theories of learning when designing such systems. For example, in *UpToDate* to facilitate easier information perception and improve its visual organization, article summary\_needs to be placed at the beginning rather than the end of the article to provide learners with an early overview of the topic. Further, to help learners visually grasp needed content, it is recommended to incorporate more visible topic headings. Finally, to overcome abundance of textual content, it is suggested to introduce

expandable/collapsible information sections to avoid lengthy textual presentation and excessive scrolling. In both *UpToDate* and *Wikipedia* it is recommended to avoid unnecessary split attention effect by providing references to images in the text (in *Wikipedia*) and making them immediately available from the same page (in *UpToDate*).

Third, designers of information retrieval systems need to provide better user awareness within the system. In *UpToDate* this could be achieved by incorporating a breadcrumb trail or other indicators (e.g., arrows to move between and within the pages) to help users understand where they are situated within the system. To avoid visual complexity of sentence-long hyperlinks in *UpToDate*, make them shorter and more concise. Additionally, to make these hyperlinks stand out against the main textual mass, change their color from dark green to blue. In *Access Medicine*, to make information acquisition more efficient, improve the loading speed of the images.

## **Future Research**

Given the importance of information literacy and self-directed learning in problem-based medical education, future studies about the appropriateness of information resources are warranted. Several research opportunities exist to build upon and strengthen the present study.

One potential study would be to explore whether individual learner characteristics had any effect on eye movement patterns. For example, fast readers demonstrate shorter fixations, longer saccades, and fewer regressions (Underwood et al., 1990). Similarly, bilingual readers make shorter fixations, longer saccades, and fewer regressions in their dominant language (Altarriba et al., 1996). Even though the student population was

considered homogenous, individual differences could have affected students' eye movement and mental effort.

Second, it has been found that grammatical and lexical text characteristics affect eye movement. For example, as word length increases, the probability of fixating (and re-fixating) on that word increases as well (Rayner & McConkie, 1976) and words of low frequency are almost always fixated (Starr & Rayner, 2001). It is also known that the number of fixations and fixation duration corresponds to the accuracy of visual inspection (Togami, 1984). In this study it was assumed that significantly higher values of total fixation duration and fixation count associated with *UpToDate* use were the result of increase in attention, not text characteristics. Additional grammatical and lexical analysis of the content (e.g., average word length, number of low frequency words, character per line) could help clarify this assumption.

Third, previous studies have found that viewing the recordings of eye movement and analyzing the eye movement patterns might stimulate learner's reflection on their own knowledge, actions, and strategies (Van Gog, Paas, & Van Merriënboer, Witte, 2005). Providing students with opportunities for cued retrospective reflection on their own information behavior may help explain the underlying motivations of the observed actions and obtain a more comprehensive picture of task performance and efficiency of learning.

Finally, future studies may include modifications of the current research. For example, the time of the task may be controlled, which may possibly affect the performance and help determine important differences related to different designs. It might also be interesting to determine whether or not students' performance changes after

each test. Finally, participants' feedback with regards to their preferences for specific design features after using all three information resources could be given more importance.

### **Conclusions**

This study presented a case of evaluating the appropriateness of three electronic information resources – *UpToDate*, *Access Medicine*, and *Wikipedia* – most commonly used for learning in pre-clerkship medical education. Findings suggest that some information resources may be less appropriate for learning based on the amount of cognitive effort they require to obtain and process needed information.

The study is distinctive in several ways. First, the research is multidisciplinary in nature and builds on the tenets of cognitive and information science. Second, the investigation relied on the eye tracking methodology, which proved to be an effective psychophysiological measure reflective of cognitive processes. This approach provided a fresh outlook on the processes of human interactions with information and information systems in the context of medical education. Third, this investigation is the first of a kind to explore the appropriateness of electronic information resources for learning in pre-clerkship medical education from the perspective of the learners' invested mental effort.

The appropriateness of electronic information resources was established in three ways: (1) by measuring and comparing the values of mental effort invested by students in interactions with each of the three resources during learning; (2) by investigating whether the experienced mental effort affected learning; and (3) by examining how certain interface design elements and information organization facilitated or interfered with information acquisition and learning. Analysis of data revealed that out of the three

resources, *UpToDate* was associated with significantly higher values of mental effort, suggesting that learners experienced overall higher cognitive load when interacting with *UpToDate*. An especially noteworthy finding was that regardless of the information resource used and the amount of invested mental effort, students were able to achieve similar learning outcomes.

The findings suggest that *UpToDate* when used alone for learning might be too demanding for medical students in pre-clerkship curriculum, especially during the first couple years of study. Despite the desire to follow the institutional trend and extensively use evidence-based resources, especially in the context of problem-based learning, students need to realize that at this stage of their education they may not yet be prepared to efficiently extract information and learn predominantly from resources like *UpToDate*. At the same time, faculty, mentors and medical librarians need to accept the fact that, despite the warnings and advisement not to, medical students continue to use questionable resources like *Wikipedia*. They also need to accept the fact that resources like *Wikipedia* might be especially beneficial for novice student learners due to their ease of use, effective structure, and information organization. As a result, it is recommended that novice learners are educated about the dangers of predominant reliance on either evidence-based or questionable resources only. Instead, students should be advised to use a combination of electronic information resources to effectively address their information needs.

## **APPENDICES**

### **Appendix A: Demographic Information Survey**

Please indicate your age range

- a. 18-21
  - b. 22-26
  - c. 27 and older
2. Please indicate your gender
- a. Male
  - b. Female
  - c. Prefer not to disclose
3. Is English your first language?
- a. Yes
  - b. No (please indicate your first language)
4. Please indicate the highest academic degree you have earned
- a. BSc
  - b. BA
  - c. MSc
  - d. MA
  - e. PhD
  - f. Professional degree (e.g., JD, MD, DO)
  - g. Other
5. Please indicate the major of the highest degree you've earned

## **Appendix B: Preferences for Electronic Information Resources Survey**

1. Please choose the types of information resource that you use most frequently for learning:
  - a. Electronic textbooks (e.g., *Harrison Online*)
  - b. Electronic textbook collections (e.g., *Access Medicine*)
  - c. Electronic textbook-like resources (e.g., *UpToDate*)
  - d. Electronic encyclopedias (e.g., *Wikipedia*)
  - e. Electronic databases (e.g., *Pubmed/Medline, Google Scholar*)
  - f. Other Internet resources (e.g., *blogs, websites*)
2. For each selected information resource you use most frequently for learning, please indicate:
  - a. the purpose of its use (e.g., for learning basic science, for finding evidence, for obtaining ideas, for confirming hypothesis, etc.)
  - b. the frequency of its use (e.g., daily, weekly, occasionally)
  - c. the reason for its use (e.g., credible, easy to use, easy to understand, recommended by the library or the faculty, etc.)
3. Please indicate your confidence in the accuracy of information available in each of the following information resources (1=not at all confident, 5 = very confident).
  - a. Electronic textbooks (e.g., *Harrison Online*)
  - b. Electronic textbook collections (e.g., *Access Medicine*)
  - c. Electronic textbook-like resources (e.g., *UpToDate*)
  - d. Electronic encyclopedias (e.g., *Wikipedia*)

- e. Electronic databases (e.g., *Pubmed/Medline, Google Scholar*)
  - f. Other Internet resources (e.g., *blogs, websites*)
4. Please indicate your confidence in the appropriateness of each of the following information resources for learning (1=not at all confident, 5 = very confident).
- a. Electronic textbooks (e.g., *Harrison Online*)
  - b. Electronic textbook collections (e.g., *Access Medicine*)
  - c. Electronic textbook-like resources (e.g., *UpToDate*)
  - d. Electronic encyclopedias (e.g., *Wikipedia*)
  - e. Electronic databases (e.g., *Pubmed/Medline, Google Scholar*)
  - f. Other Internet resources (e.g., *blogs, websites*)

## **Appendix C: Screening Test**

1. How familiar are you with Disease 1 (e.g., have you heard about it, have you independently read about/studied it, etc.)? (Likert scale, 1=not at all familiar, 5=very familiar)
2. If you indicated familiarity with Disease 1, please explain where it came from.
3. How do you evaluate your medical knowledge of the disease? (Likert scale, 1=none, 5=excellent)
4. Do you have a family member, a friend, or an acquaintance who has been diagnosed with this disease? (yes/no)
5. Do you think you are emotionally attached to this disease?

## Appendix D: Knowledge Tests

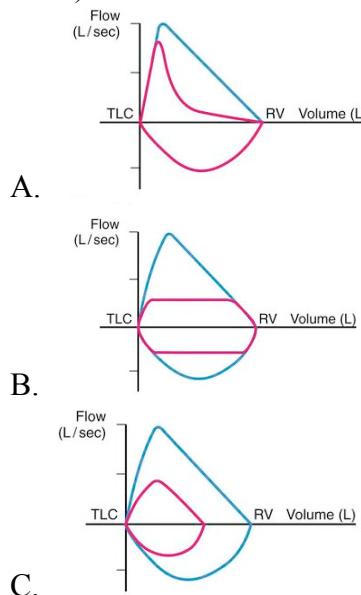
### Disease 1 - Idiopathic Pulmonary Fibrosis

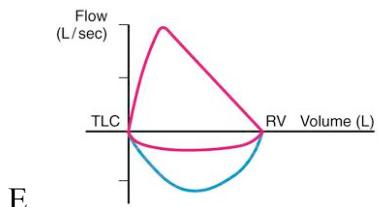
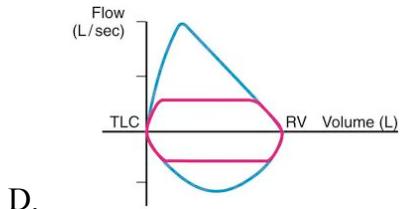
#### Primary sources used for question generation:

1. Borchers, A.T., Chang, C., Keen, C.L., & Gershwin, M.E. (2011). Idiopathic Pulmonary Fibrosis – an epidemiologic and pathologic review. *Clinical Reviews in Allergy & Immunology*, 40(2), 117-134. <http://dx.doi.org/10.1007/s12016-010-8211-5>
2. King, T.E., Pardo, A., Selman, M. (2011). Idiopathic Pulmonary Fibrosis. *Lancet*, 378(9807), 1949–1961. [http://dx.doi.org/10.1016/S0140-6736\(11\)60052-4](http://dx.doi.org/10.1016/S0140-6736(11)60052-4)
3. Steele, M.P. & Schwartz, D.A. (2013). Molecular Mechanisms in Progressive Idiopathic Pulmonary Fibrosis. *Annual Review of Medicine*, 64(1), 265-276. <http://dx.doi.org/10.1146/annurev-med-042711-142004>

#### Questions:

1. A patient is being evaluated for progressive dyspnea. He has noticed that recently when playing basketball with friends it has become harder and harder to keep up, and recently he has become too winded to even be able to play. His father and brother both have asthma, but he did not respond to a methacholine challenge test. Which of the following flow-Volume curves would most suggest a diagnosis of idiopathic interstitial pulmonary fibrosis? (Blue Line – Normal; Purple/Red Line – Patient)





- F. I don't know  
G. I can't remember reading about it

2. As a physician in a small community you become interested in a local family who has had multiple generations of family members who developed idiopathic pulmonary fibrosis. You contact pulmonologists at your state university who ask you to collect blood samples for genetic analysis. After reading up on the disease you realize that extremely rare forms seem to be linked to mutation in a gene responsible for which of the following?
- A. Cell cycle control in fibroblasts
  - B. Production of surfactant proteins
  - C. Self-tolerance of type I pneumocytes
  - D. Production of IgA
  - E. MHC I expression in alveolar macrophages
  - F. I don't know
  - G. I can't remember reading about it
3. A woman with idiopathic interstitial pulmonary fibrosis presents for increased shortness of breath and fatigue. On exam you note pitting edema, hepatomegaly, and jugular venous distension. What is the likely mechanism of the new signs in this patient?
- A. Fibrosis of lungs leads to crowding of the mediastinum, preventing complete cardiac filling in diastole.
  - B. Inflammatory cytokines return with blood from the lungs, causing damage to the left ventricle.
  - C. Increased TGF- $\beta$  signaling occurs systemically, leading to deposition of collagen in vascular walls.
  - D. Hypoxic vasoconstriction creates pulmonary hypertension and increased strain on the right ventricle.
  - E. Secondary autoimmune hepatitis

- F. I don't know
  - G. I can't remember reading about it
4. You have just diagnosed a 35-year-old patient with idiopathic pulmonary fibrosis. You caught the disease early during a routine physical in which you noted bilateral crackles, prompting a thorough pulmonary investigation. He seems distraught at the severity of the diagnosis and is in disbelief. He asks what the likely outcome of his illness will be. You explain:
- A. Only 1 in 5 patients with this diagnosis will live to 5 years.
  - B. Because you caught his disease early you can use high dose steroids to prevent deterioration.
  - C. It is likely that his disease will promote the formation of small cell lung cancer, which will be untreatable.
  - D. Newer drugs targeting fibroblast receptors can be used to extend his life considerably.
  - E. About 40% of patients live to ten years.
  - F. I don't know
  - G. I can't remember reading about it
5. A female patient is hospitalized for her lung disease. During evaluation her complete blood count shows increased hemoglobin and hematocrit with normal sized red cells. Her total white blood cell count is within normal limits. Her O<sub>2</sub> saturation is 85% on 5L of oxygen. If this patient has a diagnosis of idiopathic pulmonary fibrosis, what is the most likely cause of her O<sub>2</sub> saturation?
- A. Fibrosis reducing the surface area in the lungs for gas exchange
  - B. Fibrosis increasing the diffusion barrier to oxygen
  - C. Fibrosis leading to decreased diameter of airways, promoting lower flow rates
  - D. Air trapping causing reduced air exchange
  - E. Increased work of breathing increasing the oxygen utilization in the blood
  - F. I don't know
  - G. I can't remember reading about it
6. According to the prevailing theory for idiopathic pulmonary fibrosis, the earliest event involves which of the following?
- A. Loss of self-tolerance
  - B. Defective surfactant production
  - C. Proliferation of pulmonary fibroblasts
  - D. Damage and loss of type I pneumocytes
  - E. Increased expression of TGF-β

- F. I don't know
- G. I can't remember reading about it

7. As a 4<sup>th</sup> year medical student interested in pulmonology/critical care, you attend a conference on the subject. You enter a session on idiopathic pulmonary fibrosis and hear the speaker talking about the gradual replacement of type I pneumocytes with type II pneumocytes. You remember back to first year histology and recall this is unusual for which of the following reasons?
  - A. Type I pneumocytes are constantly regenerated like all epithelia
  - B. Type II cells are specialized for secretion of surfactant, something type I pneumocytes cannot do
  - C. Type II pneumocytes are not typically found within alveoli
  - D. Type II pneumocytes are typically non-dividing cells
  - E. Type II pneumocytes usually serve as progenitor cells for type I pneumocytes
  - F. I don't know
  - G. I can't remember reading about it
8. On teaching rounds you are assigned to a pulmonary/critical care team. As you come to one patient the instructor asks you to listen to her heart and lungs. She is notably dyspneic and you note clubbing of the fingers. On auscultation you hear diffuse inspiratory crackles. What do these crackles represent?
  - A. The opening of collapsed airways due to fibrosis
  - B. Fluid accumulation within the interstitium
  - C. Friction between fibrotic lungs and pleura
  - D. Pneumomediastinum due to alveolar rupture
  - E. Turbulent blood flow through fibrosed alveolar capillaries
  - F. I don't know
  - G. I can't remember reading about it
9. Recent advancements in treatment of idiopathic pulmonary fibrosis have seen studies experimenting with the addition of N-Acetylcysteine (NAC) to current pharmacologic regimens. What is the hypothesized mechanism for N-Acetylcysteine's proposed efficacy?
  - A. NAC stimulates surfactant release
  - B. NAC reduces the immune response by inhibition neutrophils
  - C. NAC inhibits the TGF- $\beta$  signaling cascade
  - D. NAC induces pulmonary vasodilation increasing blood flow to wash away excess TGF- $\beta$
  - E. NAC replenishes glutathione and serves as an antioxidant

- F. I don't know
- G. I can't remember reading about it

10. A patient presents to your clinic for shortness of breath and progressive fatigue with activity. On exam you hear a prominent P2 and inspiratory crackles. He is referred to radiology for a High Resolution CT. The report returns that the patient has "patchy, subpleural reticular opacities and honeycombing." Which of the following diseases should be on the top of your differential?
- A. Emphysema
  - B. Chronic bronchitis
  - C. Hypersensitivity pneumonitis
  - D. Bronchioalveolar carcinoma
  - E. Klebsiella pneumonia
  - F. I don't know
  - G. I can't remember reading about it

### **Disease 2 - Primary Sclerosing Cholangitis**

#### **Primary sources used for question generation:**

1. Chapman, R., Fevery, J., Kalloo, A., Nagorney, D.M., Boberg, K.M., Shneider, B., & Gores, G.J. (2010). Diagnosis and Management of Primary Sclerosing Cholangitis. *Hepatology*, 51(2), 660-678. <http://dx.doi.org/10.1002/hep.23294>
2. Eaton, J.E., Talkwalker, J.A., Lazaridis, K.N., Gores, G.J., & Lindor, K.D. (2013) Pathogenesis of Primary Sclerosing Cholangitis and Advances in Diagnosis and Management. *Gastroenterology*, 145, 521–536. <http://dx.doi.org/10.1053/j.gastro.2013.06.052>
3. Hirschfield, G.M., Karlsen, T.H., Lindor, D.K., & Adams, D.H. (2013). Primary Sclerosing Cholangitis. *Lancet*, 382, 1587–99. [http://dx.doi.org/10.1016/S0140-6736\(13\)60096-3](http://dx.doi.org/10.1016/S0140-6736(13)60096-3)

#### **Questions:**

1. A physician in a busy practice sees multiple patients over the course of a week. Several complain of excessive fatigue and itching. Which of the following patients is likely to have primary sclerosing cholangitis?
  - A. A 45 year old mother to 5 children
  - B. A 63 year old man with a history of colon cancer
  - C. A 35 year old male with ulcerative colitis
  - D. A 25 year old woman with type I diabetes
  - E. A 45 year old woman with autoimmune thyroiditis
  - F. I don't know

- G. I can't remember reading about it
2. A patient with a history of ulcerative colitis has a positive lab screening indicating possible primary sclerosing cholangitis. Which of the following lab values is most likely to rule in a diagnosis of primary sclerosing cholangitis?
- A. Elevated ESR
  - B. Positive anti-mitochondrial antibodies
  - C. Decreased Albumin
  - D. Elevated alkaline phosphatase
  - E. Increased AST/ALT
  - F. I don't know
  - G. I can't remember reading about it
3. A patient presents to your clinic for recurrent bouts of fever and jaundice. His wife, who is accompanying him, states that he is constantly fatigued and seems to have periods of intense itching. Initial labs demonstrate elevated liver enzymes and an elevated alkaline phosphatase. Aside from elevated bilirubin, the rest of his labs are normal. Which of the following is the next best step to evaluate this patient for a diagnosis of primary sclerosing cholangitis?
- A. Upright abdominal x-ray
  - B. MRI Cholangiography
  - C. Endoscopic retrograde cholangiopancreatography
  - D. Endoscopy
  - E. Liver biopsy
  - F. I don't know
  - G. I can't remember reading about it
4. A patient presents to a gastroenterologist with an atypical presentation of primary sclerosing cholangitis. To make the diagnosis a biopsy is performed. Which of the following lesions seen on pathology would confirm this diagnosis?
- A. Multiple Granulomas
  - B. Concentric fibrosis surrounding biliary ducts
  - C. Cytoplasmic inclusions representing rough endoplasmic reticulum
  - D. Fibrosis of the portal vein
  - E. Regenerative nodules in a fibrotic stroma
  - F. I don't know
  - G. I can't remember reading about it

5. A researcher is reviewing records and determining diagnosis bias. Of the following, which presentation is likely to be caused by primary sclerosing cholangitis?
- A 36-year-old woman with a history of bloody stools complains of excessive fatigue and itching.
  - A 50-year-old male complains of loose watery stools and an itchy rash on his elbows
  - A 15-year-old male presents with nausea/vomiting, jaundice, and discolored urine
  - A 25-year-old female with a history of lupus presents with chronic fatigue and rash
  - A 35-year-old male presents with jaundice and confusion
  - I don't know
  - I can't remember reading about it
6. A patient with advanced primary sclerosing cholangitis is hospitalized for worsening disease. Treatment is set to be initiated. Which of the following treatments has been proven to prolong life expectancy?
- IV ursodeoxycholic acid
  - High dose corticosteroids
  - Surgical grafting of artificial biliary tract
  - Colectomy of diseased segments of the colon
  - Liver transplantation
  - I don't know
  - I can't remember reading about it
7. A patient in the gastroenterology clinic has a multiyear history of primary sclerosing cholangitis. Which of the following cancers is most strongly linked to his disease?
- Cholangiocarcinoma
  - Hepatocellular carcinoma
  - Pancreatic adenocarcinoma
  - Gallbladder adenoma
  - Adenomyomatosis
  - I don't know
  - I can't remember reading about it
8. There are multiple hypotheses for the link between inflammatory bowel disease and primary sclerosing cholangitis. One plausible hypothesis is:
- Direct blood flow from portal circulation to liver allows deposition of antibody/antigen complexes

- B. Inflammation of the colon causes congestion in the portal blood flow
  - C. Bowel inflammation leads to increased absorption of bile acids, leading to biliary irritation
  - D. Both diseases are mediated by the same autoantibodies
  - E. Inversion of a single gene causes both diseases
  - F. I don't know
  - G. I can't remember reading about it
9. A female patient presents to your clinic after being hospitalized for jaundice and being diagnosed with primary sclerosing cholangitis. She has had no prior symptoms and has no other disease. Which of the following is necessary as a next step in her care?
- A. Sequence her genome to determine if she has a genetic form of disease
  - B. Place her on the transplant list
  - C. Schedule a colonoscopy with multiple biopsies
  - D. Give monoclonal antibodies against NK cells
  - E. Schedule a liver biopsy to confirm the diagnosis
  - F. I don't know
  - G. I can't remember reading about it
10. Patients with primary sclerosing cholangitis have an increased risk for hepatocellular carcinoma. This is most likely because of one of the following processes.
- A. Primary sclerosing cholangitis creates an increased risk of infection by viruses and bacteria, increasing the risk of oncogenic organism infestation.
  - B. Recurrent bouts of fibrosis and healing within the liver cause risk of metastatic change.
  - C. Regeneration of biliary epithelia causes release of growth signals, which cause also causes hepatocyte proliferation.
  - D. Ulcerative colitis increases hepatocellular carcinoma risk due to delivery of inflammatory cytokines via portal circulation – primary sclerosing cholangitis is merely another manifestation
  - E. The accumulation of bile within the biliary tract is carcinogenic to hepatocytes.
  - F. I don't know
  - G. I can't remember reading about it

### **Disease 3 - Neurosarcoidosis**

**Primary sources used for question generation:**

1. Lacomis, D. (2011) Neurosarcoidosis. *Current Neuropharmacology*, 9(3), 429-436. <http://dx.doi.org/10.2174/157015911796557975>.
2. Pawate, S., Moses, H., Sriram, S. (2009). Presentations and outcomes of neurosarcoidosis: a study of 54 cases. *QJM*, 102 (7), 449–460. <http://dx.doi.org/10.1093/qjmed/hcp042>
3. Segal, B.M. (2013). Neurosarcoidosis: diagnostic approaches and therapeutic strategies. *Current opinion in neurology*, 26(3), 307-313.

**Questions:**

1. A 30-year-old patient has gone through a lengthy diagnostic process in which he was fully worked up for multiple sclerosis. However, his final diagnosis was neurosarcoidosis. Why is it that a patient with sarcoid might be mistaken for an multiple sclerosis patient?
  - A Both multiple sclerosis and sarcoid demonstrate increased levels of ACE in the cerebrospinal fluid
  - B Neurosarcoid commonly affects the cranial nerves causing lesions such as optic neuritis
  - C The peripheral neuropathies that occurs in sarcoidosis mimic the lower motor neuron lesions found in multiple sclerosis
  - D Both are demyelinating diseases
  - E Poor diagnostic skills – these diseases actually have little overlap
  - F I don't know
  - G I can't remember reading about it
2. Which of the following patients is most likely to have neurosarcoidosis?
  - A 40-year-old African American woman with amenorrhea
  - B A 13-year-old Asian boy with polyuria
  - C An 80-year-old Caucasian woman with a history of several autoimmune diseases
  - D A 19-year-old white male with persistent dry cough
  - E A 30-year-old Latina woman who has a history of lupus and suspicious lung nodules
  - F I don't know
  - G I can't remember reading about it
3. A man presents to your clinic with a history of sarcoidosis and recent onset of neurologic dysfunction. Before you walk in to see the patient you decide to quiz the frightened medical student who has been following you around all day. You ask him what the most common cranial nerve affected by sarcoidosis is. Which of the following is the correct answer?

- A. CN V
  - B. CN VI
  - C. CN VII
  - D. CN VIII
  - E. CN IX
  - F. I don't know
  - G. I can't remember reading about it
4. While working on the neurology team you see a myriad of cases. You begin to realize that neurological diseases have a lot of overlap and often require in depth investigations to isolate a diagnosis. During the investigation of a middle-aged female patient, your attending states that neurosarcoidosis would be on the differential. Which of the following **would not** be expected as a symptom of neurosarcoidosis and would thus help you rule it down?
- A. Myopathy
  - B. C5-C6 radiculopathy
  - C. Unilateral hearing loss
  - D. Cerebrovascular accident
  - E. Internuclear ophthalmoplegia
  - F. I don't know
  - G. I can't remember reading about it
5. MRI is often used in aiding the diagnosis of neurosarcoidosis. Unfortunately, MRI images have been known to cause misdiagnosis. Which of the following is true and is likely the reason MRI is not always diagnostic for neurosarcoid?
- A. Lesions seen on MRI are not specific for neurosarcoid
  - B. Lesions seen on MRI are not sensitive for neurosarcoid
  - C. Neurosarcoid does not cause notable lesions on MRI
  - D. The required MRI cuts are beyond the range of some MRI machines
  - E. Patients with neurosarcoid often have old MRI lesions confounding the diagnosis
  - F. I don't know
  - G. I can't remember reading about it
6. During a busy week in clinic, you see a mother of 2 who has recently complained of excessive urination and thirst. She has to get up multiple times a night to urinate and the frequency has begun disrupting her day. Urinalysis reveals extremely dilute urine, prompting you to do a water deprivation test. If this woman has neurosarcoidosis related diabetes insipidus, the results of her deprivation test would be which of the following?
- A. No concentration of urine until administration of desmopressin

- B. No concentration of urine even with the administration of desmopressin
  - C. No concentration of urine until administration of demeclocycline
  - D. No concentration of urine even with the administration of demeclocycline
  - E. Normal concentration of urine with water restriction, but no increased concentration with the addition of recombinant ADH
  - F. I don't know
  - G. I can't remember reading about it
7. While shopping in Walmart, one of your rheumatology patients has a seizure that is witnessed by several bystanders. He is admitted to the inservice after presenting to the ED. This patient has a long history of sarcoidosis and you suspect this is a neurologic manifestation. Which of the following is the mechanism by which neurosarcoidosis causes seizure?
- A. Neurosarcoid often causes SIADH when the lesion occurs in the hypothalamus, the resulting electrolyte imbalances lead to seizure
  - B. Neurosarcoid lesions have a predilection for the basilar meninges, and infarction of the basilar artery can cause transient bouts of seizure
  - C. Neurosarcoid lesions are composed of infiltrating cells, which are non-depolarizing. The disruption of neuronal signaling causes over-excitation of some cells leading to seizure
  - D. Cytokines released from macrophages residing in a sarcoid lesion cross-react with glutamate receptors
  - E. Neurosarcoid lesions cause mass effect similar to meningeal tumors, creating a nidus for seizure activity
  - F. I don't know
  - G. I can't remember reading about it
8. Sarcoidosis can sometimes affect the peripheral nervous system, causing neuropathies. In cases of peripheral neurosarcoidosis, the mechanism of neuropathy is most analogous to which of the following?
- A. Guillain-Barré Syndrome
  - B. Type II Diabetes
  - C. Myasthenia Gravis
  - D. Carpal Tunnel Syndrome
  - E. Wallerian Degeneration
  - F. I don't know
  - G. I can't remember reading about it
9. A patient presents to the ED with recent onset of seizure and progressive facial weakness. On physical exam you note that she has several tattoos that have odd scarring. MRI reveals multiple small masses throughout her brain. You decide to

biopsy one of the tattoo lesions. To be consistent with neurosarcoïdosis the biopsy should display which of the following?

- A. A necrotic lesion walled off by astrocytes
- B. Migratory fibroblasts repairing damaged tissues
- C. Epithelioid cells, CD4+ T cells and a few scattered giant cells
- D. Numerous polymorphonuclear cells with few mononuclear cells
- E. A cystic lesion with peripheral fibrosis
- F. I don't know
- G. I can't remember reading about it

10. A patient presents to your office with acute onset diplopia and unilateral facial palsy. You note that the patient has a past history of sarcoidosis. The patient is anxious because she cannot see and fears permanent changes. The best initial treatment of this patient is which of the following?

- A. Surgery to remove the granulomatous mass before permanent damage occurs
- B. Corticosteroids to reduce current inflammation
- C. Vasodilators to allow collateral blood flow alleviating infarcted areas
- D. Draw labs to investigate a possible infection triggering this exacerbation
- E. Anxiolytic drugs to calm the patient as there is no actual treatment for neurosarcoïdosis
- F. I don't know
- G. I can't remember reading about it

## **Appendix E: Interview Protocol**

1. Please tell me about your experience learning about diseases in different information resources?
  - a. Which disease was easier and more difficult to read about and understand? Why do you think so?
  - b. What did you think about the format and the presentation of the material (e.g., font, space between the lines, text density, etc.)?
  - c. What did you think about the complexity of the material presentation (e.g., unfamiliar words/terminology, sentence structures, reading difficulty)?
  - d. What did you think about the presence of images?
  - e. What did you think about the presence of hyperlinks?
  - f. Based on your answers, which do you think is more appropriate for learning about basic science?
2. Which elements/characteristics of these information resources do you think contributed to/interfered with your learning?
3. Why do/don't you like to use these information resources?
4. Which of the information resources did you think was easier/more difficult to use? Why?
5. Which of the information resources do you think provided the most accurate information? Why?
6. How do you think you did on the comprehension test? How did you use prior knowledge?

7. What do you think could have affected your performance, e.g., the content was difficult to understand?
8. Did your attitude towards/confidence in the information resource change after the test? Why/why not?
9. How much time do you usually spend learning new material from an information resource?
10. How do you usually interact with these information resources during learning, e.g., do you take notes, write summaries, print content, etc.?

## REFERENCES

- Ahmadi, S.-F., Faghankhani, M., Javanbakht, A., Akbarshahi, M., Mirghorbani, M., Safarnejad, B., & Baradaran, H. (2011). A comparison of answer retrieval through four evidence-based textbooks (ACP PIER, Essential Evidence Plus, First Consult, and UpToDate): A randomized controlled trial. *Medical Teacher*, 33(9), 724–730. doi:10.3109/0142159X.2010.531155
- Allahwala, U. K., Nadkarni, A., & Sebaratnam, D. F. (2013). Wikipedia use amongst medical students - new insights into the digital revolution. *Medical Teacher*, 35(4), 337. doi:10.3109/0142159X.2012.737064
- Altarriba, J., Kroll, J., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory & Cognition*, 24(4), 477–492.
- Amadieu, F., Tricot, A., & Mariné, C. (2009). Prior knowledge in learning from a non-linear electronic document: Disorientation and coherence of the reading sequences. *Computers in Human Behavior*, 25(2), 381–388. doi:10.1016/j.chb.2008.12.017
- Antonenko, P. D., & Niederhauser, D. S. (2010). The influence of leads on cognitive load and learning in a hypertext environment. *Computers in Human Behavior*, 26(2), 140–150. doi:10.1016/j.chb.2009.10.014
- Azer, S. A. (2014). Evaluation of gastroenterology and hepatology articles on Wikipedia: are they suitable as learning resources for medical students? *European Journal of Gastroenterology & Hepatology*, 26(2), 155–163. doi:10.1097/MEG.0000000000000003
- Azevedo, R. (2005). Using hypermedia as a metacognitive tool for enhancing student learning? The role of self-regulated learning. *Educational Psychologist*, 40, 199–209.
- Baddeley, A. (1992). Working memory. *Science*, 255, 556–559.
- Barrows, H. S. (2000). *Problem-Based Learning Applied to Medical Education*. Southern Illinois University Press, Springfield.
- Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, 91(2), 276–292.
- Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Bernston (Eds.), *Handbook of Psychophysiology* (2nd ed., pp. 142–162). Cambridge University Press.

- Betrancourt, M. (2005). The animation and interactivity principles in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 287–296). Cambridge MA: Cambridge University Press.
- Bhattacharyya, D., Chowdhury, B., Chatterjee, T., Pal, M., & Majumdar, D. (2014). Selection of character/background colour combinations for onscreen searching tasks: An eye movement, subjective and performance approach. *Displays*, 35(3), 101–109. doi:10.1016/j.displa.2014.03.002
- Blumberg, P., & Michael, J. A. (1992). Development of self-directed learning behaviors in a partially teacher-directed problem-based learning curriculum. *Teaching and Learning in Medicine*, 4(1), 3–8.
- Boumarafi, B. (2010). Electronic resources at the University of Sharjah medical library: an investigation of students' information-seeking behavior. *Medical Reference Services Quarterly*, 29(4), 349–62. doi:10.1080/02763869.2010.518921
- Bruneau, D., Sasse, M. A., & McCarthy, J. D. (2002). The eyes never lie: The use of eye tracking data in HCI research. In *Proceedings of the CHI'02 Workshop on Physiological Computing*. NY: ACM Press.
- Brünken, R., Plass, J. L., & Leutner, D. (2010). Direct Measurement of Cognitive Load in Multimedia Learning. *Educational Psychologist*, 38(1), 53–61. doi:10.1207/S15326985EP3801
- Brünken, R., Steinbacher, S., Plass, J. L., & Leutner, D. (2002). Assessment of Cognitive Load in Multimedia Learning Using Dual-Task Methodology. *Experimental Psychology*, 49(2), 109–119. doi:10.1027//1618-3169.49.2.109
- Campbell, R., & Ash, J. (2006). An evaluation of five bedside information products using a user-centered, task-oriented approach. *Journal of Medical Library Association*, 94(4), 435–441, e206–e207.
- Carver, R. P. (1971). Pupil dilation and its relationship to information processing during reading and listening. *The Journal of Applied Psychology*, 55(2), 126–34. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/5551633>
- Case, S. M., & Swanson, D. B. (2002). *Constructing written test questions for the basic and clinical sciences* (3rd ed.). Philadelphia, PA: National Board of Medical Examiners. Chicago. Retrieved from <http://www.nbme.org/publications/item-writing-manual-download.html>
- Chen, K. -n., Lin, P. -c., Chang, S.-S., & Sun, H. -c. (2011). Library use by medical students: A comparison of two curricula. *Journal of Librarianship and Information Science*, 43(3), 176–184. doi:10.1177/0961000611410928

- Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. Sternberg (Ed.), *Advances in the Psychology of Human Intelligence* (pp. 7–76). Hillsdale, NJ: Erlbaum.
- Cierniak, G., Scheiter, K., & Gerjets, P. (2009). Explaining the split-attention effect: Is the reduction of extraneous cognitive load accompanied by an increase in germane cognitive load? *Computers in Human Behavior*, 25(2), 315–324. doi:10.1016/j.chb.2008.12.020
- Cogdill, K. W., & Moore, M. E. (1997). First-year medical students' information needs and resource selection: responses to a clinical scenario. *Bulletin of the Medical Library Association*, 85(1), 51–54. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=226223&tool=pmcentrez&rendertype=abstract>
- Conklin, J. (1987). Hypertext : An Introduction and Survey. *Computer*, 20(9), 17–41.
- Cooper, A. L., & Elnicki, D. M. (2011). Resource utilisation patterns of third-year medical students. *The Clinical Teacher*, 8(1), 43–7. doi:10.1111/j.1743-498X.2010.00393.x
- Crawford, H. (2001). Encyclopedias. In L. C. S. R. Bopp (Ed.), *Reference and information services: an introduction* (3 ed., pp. 433–459). Englewood, CO: Libraries Unlimited.
- de Jong, T. (2010). Cognitive load theory, educational research, and instructional design: Some food for thought. *Instructional Science*, 38(2), 105–134. doi:10.1007/s11251-009-9110-0
- DeLeeuw, K. E., & Mayer, R. E. (2008). A comparison of three measures of cognitive load: Evidence for separable measures of intrinsic, extraneous, and germane load. *Journal of Educational Psychology*, 100, 223–234.
- Deretchin, L. F., Yeoman, L. C., & Seidel, C. L. (1998). Student Information Resource Utilization in Problem-Based Learning Baylor College of Medicine, 1–6.
- DeStefano, D., & LeFevre, J. A. (2007). Cognitive load in hypertext reading: A review. *Computers in Human Behavior*, 23(3), 1616–1641. doi:10.1016/j.chb.2005.08.012
- Dillon, A., & Jobst, J. (2005). Multimedia learning with hypermedia. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 569–588). Cambridge MA: Cambridge University Press.
- Dodd, L. (2007). The Impact of Problem-Based Learning on the Information Behavior and Literacy of Veterinary Medicine Students at University College Dublin. *The*

*Journal of Academic Librarianship*, 33(2), 206–216.  
doi:10.1016/j.acalib.2006.12.008

Duchowski, A. T. (2002). A breadth-first survey of eye-tracking applications. *Behavior Research Methods, Instruments, & Computers*, 34(4), 455–470. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12564550>

Duchowski, A. T. (2003). *Eye tracking methodology: Theory and practice*. London: Springer-Verlag Ltd.

Erdelez, S. (1999). Information encountering: It's more than just bumping into information. *Bulletin of the American Society for Information Science*, 25(3), 25–29.

Eskola, E.-L. (2005). Information literacy of medical students studying in the problem-based and traditional curriculum. *Information Search*, 10(2).

Esparza, J. M., Shi, R., McLarty, J., Comegys, M., & Banks, D. E. (2013). The effect of a clinical medical librarian on in-patient care outcomes. *Journal of the Medical Library Association: JMLA*, 101(3), 185–91. doi:10.3163/1536-5050.101.3.007

Fallis, D. (2008). Toward an Epistemology of Wikipedia. *Journal of the American Society for Information Science and Technology*, 59(May), 1662–1674.  
doi:10.1002/asi

Fenton, S. H., & Badgett, R. G. (2007). A comparison of primary care information content in UpToDate and the National Guideline Clearinghouse. *Journal of the Medical Library Association*, 95(3), 255–259. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1924927/>

Fitts, P. M., Jones, R. E., & Milton, J. L. (2005). Eye movement of aircraft pilots during instrument-landing approaches. In *Ergonomics: Psychological Mechanisms and Models in Ergonomics* (pp. 56–66). New York: Taylor & Francis.

Foltz, P. W. (1996). Comprehension, coherence, and strategies in hypertext and linear text. In J.-F. Rouet, J. J. Levonen, A. Dillon, & R. J. Spiro (Eds.), *Hypertext and Cognition* (pp. 109– 136). Mahwah, NJ: Lawrence Erlbaum Associates.

Gaitsgory, O., Burgess, A., & Mellis, C. (2013). Opinion piece: “Medical students - learning from textbooks or electronic media?” *Journal of Paediatrics and Child Health*, 49(9), E370–E372. doi:10.1111/jpc.12180

Gall, J. E., & Hannafin, M. J. (1994). A framework for the study of hypertext. *Instructional Science*, 22(207-232).

- Geissler, G., Zinkhan, G., & Watson, R. (2001). Web home page complexity and communication effectiveness. *Journal of the Association for Information Systems*, 2(2).
- Gerjets, P., & Scheiter, K. (2003). Goal configurations and processing strategies as moderators between instructional design and cognitive load : Evidence from hypertext-based instruction. *Educational Psychologist*, 38(1), 33–41.
- Gerjets, P., Scheiter, K., Opfermann, M., Hesse, F. W., & Eysink, T. H. S. (2009). Learning with hypermedia: The influence of representational formats and different levels of learner control on performance and learning behavior. *Computers in Human Behavior*, 25(2), 360–370. doi:10.1016/j.chb.2008.12.015
- Goldberg, H. J., & Wichansky, A. M. (2002). Eye Tracking in Usability Evaluation: A Practitioner's Guide. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The Mind's Eyes: Cognitive and Applied Aspects of Eye Movements* (pp. 493–516). Elsevier Science, Oxford.
- Goldberg, J. H. (2012). Relating Perceived Web Page Complexity to Emotional Valence and Eye Movement Metrics Impression of Complexity. In *Proceedings of the Human Factors and Ergonomics Society 56th Annual Meeting* (pp. 501–505).
- Guyatt, G., Meade, M. O., Rennie, D., & Cook, D. J. C. (2008). *Users' Guides to the Medical Literature* (2nd ed.). McGraw-Hill Medical. doi:10.1036/007159034X
- Haider, H., & Frensch, P. A. (1999). Eye Movement During Skill Acquisition : More Evidence for the Information-Reduction Hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(1), 172–190.
- Haigh, C. (2011). Wikipedia as an evidence source for nursing and healthcare students. *Nurse Education Today*, 31(2), 135–9. doi:10.1016/j.nedt.2010.05.004
- Harper, S., Michailidou, E., & Stevens, R. (2009). Toward a definition of visual complexity as an implicit measure of cognitive load. *ACM Transactions on Applied Perception*, 6(2), 1–18. doi:10.1145/1498700.1498704
- Hayward, R. S., El-Hajj, M., Voth, T. K., & Deis, K. (2006). Patterns of use of decision support tools by clinicians. In *AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium* (pp. 329–333).
- Head, A. J., & Eisenberg, M. B. (2010). How today's college students use Wikipedia for course-related research. *First Monday*, 15(3). Retrieved from <http://uncommonculture.org/ojs/index.php/fm/article/view/2830>
- Heller, R. (1990). The role of hypermedia in education: A look at the research issues. *Journal of Research on Computing in Education*, 22(4), 431–441.

- Heyd, M. (2010). Three e-Book Aggregators for Medical Libraries: NetLibrary, Rittenhouse R2 Digital Library, and STAT!Ref. *Journal of Electronic Resources in Medical Libraries*, 7(1), 13–41. doi:10.1080/15424060903585693
- Hilligoss, B., & Rieh, S. Y. (2008). Developing a unifying framework of credibility assessment: Concept, heuristics, and interaction in context. *Information Processing & Management*, 44, 1467–1484.
- Hoffman, S. (1997). Elaboration theory and hypermedia: Is there a link? *Educational Technology*, 37(1), 57–64.
- Holsanova, J., Holmberg, N., & Holmqvist, K. (2008). Reading information graphics: The role of spatial contiguity and dual attentional guidance. *Applied Cognitive Psychology*, 22, 1–12. doi:10.1002/acp
- Hughes, B., Joshi, I., Lemonde, H., & Wareham, J. (2009). Junior physician's use of Web 2.0 for information seeking and medical education: a qualitative study. *International Journal of Medical Informatics*, 78(10), 645–55. doi:10.1016/j.ijmedinf.2009.04.008
- Hyönä, J. (2010). The use of eye movements in the study of multimedia learning. *Learning and Instruction*, 20(2), 172–176. doi:10.1016/j.learninstruc.2009.02.013
- Hyönä, J., & Lorch, R. F. (2004). Effects of topic headings on text processing: evidence from adult readers' eye fixation patterns. *Learning and Instruction*, 14(2), 131–152. doi:10.1016/j.learninstruc.2004.01.001
- Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during fixations in reading: effects of word frequency. *Attention, Perception, & Psychophysics*, 40(6), 431–439.
- Iqbal, S. T., Zheng, X. S., & Bailey, B. P. (2004). Task-evoked pupillary response to mental workload in human-computer interaction. In *Proceedings of the 2004 conference on Human factors and computing systems - CHI '04* (p. 1477). New York, New York, USA: ACM Press. doi:10.1145/985921.986094
- Jacobson, J. Z., & Dodwell, P. C. (1979). Saccadic eye movements during reading. *Brain and Language*, 8(3), 303–314.
- Jacobson, M. J., & Spiro, R. J. (1995). Hypertext learning environments, cognitive flexibility, and the transfer of complex knowledge: an empirical investigation. *Journal of Educational Computing Research*, 12(4), 301–333.
- Jaeggi, S. M., Buschkuhl, M., Etienne, A., Ozdoba, C., Perrig, W. J., & Nirkko, A. C. (2007). On how high behavioral performers keep cool brains in situations of cognitive overload. *Cognitive Affective & Neuroscience*, 7, 75–89.

- Jonassen, D. H. (1999). Designing constructivist learning environments. In *Instructional design theories and models: A new paradigm of instructional theory* (pp. 215–239).
- Jonassen, D. H., & Grabinger, R. S. (1990). Problems and issues in designing hypertext/hypermedia for learning. In D. H. Jonassen & H. Mandl (Eds.), *Designing Hypermedia for Learning* (pp. 3–25). Berlin: Springer.
- Judd, T., & Kennedy, G. (2011). Expediency-based practice? Medical students' reliance on Google and Wikipedia for biomedical inquiries. *British Journal of Educational Technology*, 42(2), 351–360. doi:10.1111/j.1467-8535.2009.01019.x
- Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, 38(1), 23–31. doi:10.1207/S15326985EP3801\_4
- Kellar, M., Watters, C., & Shepherd, M. (2007). A field study characterizing Web-based information-seeking tasks. *Journal of the American Society for Information Science & Technology*, 58(7), 999–1018.
- Kintsch, W. (1998). *Comprehension—A paradigm for cognition*. Cambridge: Cambridge University Press.
- Kolers, P. A., Duchnicky, R. L., & Ferguson, D. C. (1981). Eye movement measurement of readability of CRT displays. *Human Factors*, 23, 517–552.
- Kommers, P. A. M., Grabinger, S., & Dunlap, J. C. (Eds.). (1996). *Hypermedia Learning Environments: Instructional Design and Interaction*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc., Publishers.
- Kuhl, J. (1984). Volitional aspects of achievement motivation and learned helplessness: Toward a comprehensive theory of action control. In B. A. Maher & W. B. Maher (Eds.), *Progress in Experimental Personality Research* (pp. 99–171). Academic Press, Inc.
- Landoni, M. (2003). Electronic books. In J. Feather & P. Sturges (Eds.), *International Encyclopedia of Information and Library Science* (2nd ed., pp. 168–171). Routledge, London Lee.
- Lavsa, S. M., Corman, S. L., Culley, C. M., & Pummer, T. L. (2011). Reliability of Wikipedia as a medication information source for pharmacy students. *Currents in Pharmacy Teaching and Learning*, 3(2), 154–158.
- Lawrence, S. L., Grosenick, D. J., Simpson, D. E., & Van Susteren, T. J. (1992). A comparison of problem-based and didactic approaches to learning on an ambulatory medicine clerkship. 4:221-4. *Teaching and Learning in Medicine*, 221–224.

Leithner, A., Maurer-Ertl, M., Glehr, M., Friesenbichler, J., Leithner, K., & Windhager, R. (2010). Wikipedia and osteosarcoma: a trustworthy patients' information? *Journal of American Medical Association*, 303(17), 373–374. doi:10.1136/jamia.2010.004507

Leuthold, S., Schmutz, P., Bargas-Avila, J. A., Tuch, A. N., & Opwis, K. (2011). Vertical versus dynamic menus on the world wide web: Eye tracking study measuring the influence of menu design and task complexity on user performance and subjective preference. *Computers in Human Behavior*, 27, 459–472. doi:10.1016/j.chb.2010.09.009

Lim, S. (2009). How and Why Do College Students Use Wikipedia? *Journal of the American Society for Information Science*, 60(11), 2189–2202. doi:10.1002/asi

Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Sciences*, 4(1), 6–14. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10637617>

Lorch Jr., R. F., & Lorch, E. P. (1996). Effects of headings on text recall and summarization. *Contemporary Educational Psychology*, 21, 261–278.

Lorch Jr., R. F., Lorch, E. P., & Inman, W. E. (1993). Effects of signaling topic structure on text recall. *Journal of Educational Psychology*, 85, 281–290.

Lorch, R. F. (1993). Integration of topic and subordinate information during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(5), 1071–1081.

Madrid, R. I., Van Oostendorp, H., & Puerta Melguizo, M. C. (2009). The effects of the number of links and navigation support on cognitive load and learning with hypertext: The mediating role of reading order. *Computers in Human Behavior*, 25(1), 66–75. doi:10.1016/j.chb.2008.06.005

Mayer, R. E. (2003). *Learning and Instruction*. Upper Saddle River, NJ: Prentice Hall.

Mayer, R. E. (2005). Cognitive Theory of Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 31–48). Cambridge MA: Cambridge University Press.

Mayer, R. E., & Chandler, P. (2001). When learning is just a click away: Does simple interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology*, 93(2), 390–397.

Mayer, R. E., Dow, G. T., & Mayer, S. (2003). Multimedia learning in an interactive self-explaining environment: What works in the design of agent-based microworlds? *Journal of Educational Psychology*, 95(4), 806–812.

Mayer, R. E., & Moreno, R. (2003). Nine Ways to Reduce Cognitive Load in Multimedia Learning. *Educational Psychologist*, 38(1), 43–52.  
doi:10.1207/S15326985EP3801\_6

Metalis, S. A., Rhoades, B. K., Hess, E. H., & Petrovich, S. B. (1980). Pupillometric assessment of reading using materials in normal and reversed orientations. *The Journal of Applied Psychology*, 65(3), 359–63. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/7380778>

Meyer, R. E. (1999). Designing instruction for constructivist learning. In *Instructional-design theories and models: A new paradigm of instructional theory* (pp. 141–159).

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

Miniukovich, A., & De Angeli, A. (2014). Quantification of interface visual complexity. In *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces* (pp. 153–160). ACM.

Montgomery, D. C. (2001). *Design and Analysis of Engineering Experiments* (5th ed.). New York: John Wiley & Sons.

Morrison, R. E., & Inhoff, A. W. (1981). Visual factors and eye movements in reading. *Visible Language*, 15, 129–146.

Natrella, M. (2010). *NIST/SEMATECH e-handbook of Statistical Methods*.

Niederhauser, D. S. (2007). Educational hypertext. In *Handbook of Research on Educational Communications and Technology: A Project of the Association for Educational Communications and Technology* (pp. 199–210).

Nielsen, J. (1995). 10 Usability Heuristics for User Interface Design. *Nielsen Normal Group*. Retrieved from <http://www.nngroup.com/articles/ten-usability-heuristics/>

Norman, D. A. (1982). *Learning and Memory*. New York: W.H. Freeman.

Paas, F., Ayres, P., & Pachman, M. (2008). Assessment of Cognitive Load in Multimedia Learning: Theory, Methods, and Applications. In D. H. Robinson & G. Schraw (Eds.), *Recent Innovations in Educational Technology that Facilitate Student Learning* (pp. 11–35). Information Age Publishing.

Paas, F., Ayres, P., & Pachman, M. (2008). Assessment of Cognitive Load in Multimedia Learning. Theory, Methods, and Applications. In D. H. Robinson & G. Schraw (Eds.), *Recent Innovations in Educational Technology that Facilitate Student Learning* (pp. 11–35). Information Age Publishing.

- Paas, F., & Van Merriënboer, J. J. G. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review*, 6(4), 351–371.
- Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: recent developments. *Educational Psychologist*, 38(1), 1–4.
- Paas, F., Renkl, A., & Sweller, J. (2004). Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. *Instructional Science*, 32, 1–8.
- Paas, F., Van Gerven, P. W., & Wouters, P. (2007). Instructional efficiency of animation: Effects of interactivity through mental reconstruction of static key frames. *Applied Cognitive Psychology*, 21(6), 783–793.
- Paas, F., & van Merriënboer, J. J. G. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review*, 6, 51–71.
- Partala, T., & Surakka, V. (2003). Pupil size variation as an indication of affective processing. *International Journal of Human-Computer Studies*, 59(1-2), 185–198. doi:10.1016/S1071-5819(03)00017-X
- Paterson, K. B., Liversedge, S. P., & Underwood, G. (1999). The influence of focus operators on syntactic processing of short relative clause sentences. *Quarterly Journal of Experimental Psychology*, 52A, 717–737.
- Pender, M. P., Lassere, B. A., Kruesi, L., Del Mar, C., & Anuradha, S. (2008). Putting Wikipedia to the Test : A Case Study. In *Special Libraries Association Annual Conference*. Seattle.
- Peterson, M. W., Rowat, J., Kreiter, C., & Mandel, J. (2004). Medical students' use of information resources: is the digital age dawning? *Academic Medicine : Journal of the Association of American Medical Colleges*, 79(1), 89–95. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/14691004>
- Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. *Learning and Instruction*, 12, 61–86.
- Poock, G. K. (1973). Information processing vs pupil diameter, 1000–1002.
- Poole, A., & Ball, L. J. (2005). Eye Tracking in Human-Computer Interaction and Usability Research : Current Status and Future Prospects. In C. Ghaoui (Ed.), *Encyclopedia of Human-Computer Interaction*. Pennsylvania: Idea Group, Inc.
- Potelle, H., & Rouet, J. F. (2003). Effects of content representation and readers' prior knowledge on the comprehension of hypertext. *International Journal of Human-Computer Studies*, 58(3), 327–345.

- Protopsaltis, A., & Bouki, V. (2006). The effects of reading goals in hypertext reading. In *Proceedings of the 24th annual ACM international conference on Design of communication* (pp. 29–34). ACM.
- Rankin, J. a. (1992). Problem-based medical education: effect on library use. *Bulletin of the Medical Library Association*, 80(1), 36–43. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1383312/>
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372–422. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1383312/>
- Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, 62(8), 1457–1506. doi:10.1080/17470210902816461
- Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior*, 22(3), 358–374.
- Rayner, K., & McConkie, G. W. (1976). What guides a reader's eye movement? *Vision Research*, 16, 829–837.
- Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: a comparison of two types of models. *Journal of Experimental Psychology*, 22(5), 1188–1200.
- Reigeluth, C. M., & Stein, F. S. (1983). The Elaboration Theory of Instruction. In C. M. Reigeluth (Ed.), *Instructional Design Theories and Models: An Overview of their Current Status*. Lawrence Erlbaum Associates, Inc., Publishers.
- Replicated Latin Squares, STAT 503 Design of Experiments. (2015). *The Pennsylvania State University Online Courses*. Retrieved from <https://onlinecourses.science.psu.edu/stat503/node/22>
- Roberts, M. J., & Newton, E. J. (2001). Understanding strategy selection. *International Journal of Human-Computer Studies*, 54(1), 137–154. doi:10.1006/ijhc.2000.0434
- Salmeron, L., Canas, J. J., Kintsch, W., & Fajardo, I. (2005). Reading strategies and hypertext comprehension. *Discourse Processes*, 40(3), 171e191.
- Salmeron, L., Kintsch, W., & Canas, J. (2006). Reading strategies and prior knowledge in learning from hypertext. *Memory & Cognition*, 34(5), 1157–1171.

- Salmerón, L., Kintsch, W., & Kintsch, E. (2010). Self-regulation and link selection strategies in hypertext. *Discourse Processes*, 47(3), 175–211.
- Savery, J. R. (2006). Overview of Problem-based Learning : Definitions and Distinctions. *The Interdisciplinary Journal of Problem-Based Learning*, 1(1), 9–20.
- Scheiter, K., & Gerjets, P. (2007). Learner Control in Hypermedia Environments. *Educational Psychology Review*, 19(3), 285–307. doi:10.1007/s10648-007-9046-3
- Schwartz, N. H., Andersen, C., Hong, N., Howard, B., & McGee, S. (2004). The influence of metacognitive skills on learners' memory of information in a hypermedia environment. *Journal of Educational Computing Research*, 31(1), 77–93.
- Smith, E. E., & Jonides, J. (1997). Working memory: A view from neuroimaging. *Cognitive Psychology*, 33(1), 5–42.
- Snow, R. E. (1980). Aptitude, learner control, and adaptive instruction. *Educational Psychologist*, 15, 151– 158.
- Spiro, R. J., Coulson, R. L., Feltovich, P. J., & Anderson, D. K. (1988). *Cognitive flexibility theory: Advanced knowledge acquisition in ill-structured domains* (Tech. Rep. No. 441). Urbana-Champaign, IL. Retrieved from <http://eric.ed.gov/?id=ED302821>
- Starr, M. S., & Rayner, K. (2001). Eye movements during reading: some current controversies. *Trends in Cognitive Sciences*, 5(4), 156–163. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11287269>
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285. doi:10.1016/0364-0213(88)90023-7
- Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, 4, 295–312.
- Sweller, J., Merrienboer, J. J. G. Van, & Paas, F. G. W. C. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review*, 10(3), 251–296.
- Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review*, 10(3), 251–296.
- Tao, D. (2008). Understanding intention to use electronic information resources: a theoretical extension of the technology acceptance model (TAM). In *Annual Symposium Proceedings of the American Medical Informatics Association* (p. 717).
- Tobii. (n.d.).

- Togami, H. (1984). Effects on visual search performance of individual differences in fixation time and number of fixations. *Ergonomics*, 27, 789 – 799.
- Underwood, G., Chapman, P., Brocklehurst, N., Underwood, J., & Crundall, D. (2003). Visual attention while driving: Sequences of eye fixations made by experienced and novice drivers. *Ergonomics*, 46, 629–646.
- Underwood, G., Hubbard, A., & Wilkinson, H. (1990). Eye fixations predict reading comprehension: The relationships between reading skill, reading speed, and visual inspection. *Language and Speech*, 33(1), 69–81.
- User Manual — Tobii Studio*. (2012) (Version 3.). Tobii Technology AB.
- Uttal, W. R., & Smith, P. (1968). Recognition of alphabetic characters during voluntary eye movements. *Perception & Psychophysics*, 3, 257–264.
- Van Gog, T., Paas, F., Van Merriënboer, J. J. G., & Witte, P. (2005). Uncovering the problem-solving process: Cued retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology: Applied*, 11, 237e244.
- Van Gog, T., & Scheiter, K. (2010). Eye tracking as a tool to study and enhance multimedia learning. *Learning and Instruction*, 20(2), 95–99.  
doi:10.1016/j.learninstruc.2009.02.009
- Van Merriënboer, J. J. G., & Sweller, J. (2010). Cognitive load theory in health professional education: design principles and strategies. *Medical Education*, 44(1), 85–93. doi:10.1111/j.1365-2923.2009.03498.x
- Verwhey, W. B., & Veltman, H. A. (1996). Detecting short periods of elevated workload: A comparison of nine workload assessment techniques. *Journal of Experimental Psychology: Applied*, 2(3), 270–285.
- Wästlund, E., Norlander, T., & Archer, T. (2008). The effect of page layout on mental workload: A dual-task experiment. *Computers in Human Behavior*, 24(3), 1229–1245. doi:10.1016/j.chb.2007.05.001
- Winne, P. H. (1995). Inherent details in self-regulated learning. *Educational Psychologist*, 30, 173–187.
- Winne, P. H. (2001). Self-regulated learning viewed from models of information processing. In B. Zimmerman & D. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (pp. 155–189). Mahwah, NJ: Erlbaum.
- Wittrock, M. C. (1990). Generative model of comprehension. *Educational Psychologist*, 24(4), 345–376.

- Wood, A., & Struthers, K. (2010). Pathology education, Wikipedia and the Net generation. *Medical Teacher*, 32(7), 618. doi:10.3109/0142159X.2010.497719
- Wood, A., & Struthers, K. (2010). Pathology education, Wikipedia and the Net generation. *Medical Teacher*, 32(7), 618. doi:10.3109/0142159X.2010.497719
- Woods, D. R. (1994). *Problem-based learning: how to gain the most from PBL*. Hamilton: McMaster University.
- Xie, B., & Salvendy, G. (2000). Prediction of mental workload in single and multiple tasks environments. *International Journal of Cognitive Ergonomics*, 4(3), 213–242.
- Zumbach, J. (2006). Cognitive overhead in hypertext learning reexamined : Overcoming the myths. *Journal of Educational Multimedia and Hypermedia*, 15(4), 411–432.
- Zumbach, J., & Mohraz, M. (2008). Cognitive load in hypermedia reading comprehension: Influence of text type and linearity. *Computers in Human Behavior*, 24(3), 875–887. doi:10.1016/j.chb.2007.02.015

## **VITA**

Dinara Saparova was born in Almaty, Kazakhstan. She received a B.A. in Informatics and English Language from Almaty State University in 2000 and a M.A. in Comparative Linguistics from Kazakh University of International Relations and World Languages in 2002. Before she enrolled in the School of Information Science & Learning Technologies at the University of Missouri (MU) to pursue her doctoral degree, she worked as a senior lecturer at the Department of Continuing and Distance Education at the Kazakh University of International Relations and World Languages from 2002 to 2009. During the doctoral program, Dinara worked as a research assistant at the MU Information Experience Laboratory and the Dean's Office at the MU School of Medicine and a teaching assistant at the MU School of Information Science and Learning Technologies, Department of Health Management and Informatics, and the School of Health Professions. Dinara is a recipient of Fulbright Grant from the Council for International Exchange of Scholars, USA, and two graduate scholarships (Phi Delta Kappa and Charles S. and Theda H. Estep) from the MU College of Education. In 2014 she was nominated for a Graduate Scholar of the Year from the MU College of Education.