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

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

By
NATHAN J. LOWRANCE
Dr. Heather Lea Moulaison, Dissertation Supervisor
December 2015
The undersigned, appointed by the dean of the Graduate School, have examined the dissertation entitled:

GOOGLE SCHOLAR AND META DESCRIPTIONS: DOES ADDING ABSTRACTS TO A SEARCH ENGINE RESULTS PAGE AID IN UNDERGRADUATE DOCUMENT TRIAGE EFFICIENCY?

Presented by Nathan J. Lowrance
A candidate for the degree of Doctor of Philosophy

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

___________________________________________
Dr. Heather Lea Moulaison

___________________________________________
Dr. Joi Moore

___________________________________________
Dr. Jenny Bossaller

___________________________________________
Dr. Nicole Campione-Barr
Dedication

This work is dedicated to my loving parents, John W. Lowrance and Heidi A. Lowrance.

Thanks for giving me the tools and support I needed, and for taking me out into the world.
Acknowledgements

I wish to express appreciation for all the efforts of my advisor Heather Lea Moulaison. I would also like to recognize the members of my research committee – Professors Joi Moore, Jenny Bossaller, and Nicole Champione-Barr. Their mentorship and support have been a most helpful blessing, and any development of my own comes only because I have the support and dedication of these fine women. I’d also like to thank Jeffery Belden and Richelle Koopman for my initiation into the world of electronic health records and Dr. John Budd for his guidance though my master’s program.

I wish all the best to my friends and colleagues I met at the University of Missouri, past and present, and would like to thank them for their conversations and friendship: Kristine Stewart, Ben Richardson, Josipa Basic, Dinara Saparova, Sean Burns, Kyle Denninger, Neely Current, A.J. Million, Keith Haggardly, Kodjo Atisom, Borchuluun Yadamsuren and Yanfei Ma. Also I’d like to give a special thanks to Douglas Abrams, this study’s subject matter expert, for his time and expertise.

I also have to thank my friend Jeff Ousley, for letting me stay with him over the summer of 2015, and Mike Bates and Robert Presson for their constant support. A special thanks goes to my academic bro-mate and workout partner Christopher Larson. These last four years I could often be found at the gym in high end intellectual debate with Chris, that or in the field practicing warrior skills, fulfilling a childhood dream, thanks to his One Shepherd program. To my friend and fellow academic Kathleen Turner, this dissertation’s literature review is dedicated to you.

This acknowledgement would not be complete without an offering of gratitude and respect to my father John Lowrance, for his tireless support and editing assistance, without which I would have struggled. Thanks and best wishes to all.
# Table of Contents

Acknowledgements ................................................................. ii

Table of Figures ................................................................. viii

Abstract .............................................................................. x

Chapter 1. INTRODUCTION ...................................................... 1

Statement of Problem ............................................................. 1

Statement of Purpose ............................................................. 3

SERP Meta Descriptions and Document Triage ....................... 4

Significance of the Research ..................................................... 6

Experimental Design and Rational .......................................... 7

Research Questions ................................................................ 7

Expected Outcomes ............................................................... 9

Conclusion ........................................................................... 9

Definitions ........................................................................... 10

Chapter 2. LITERATURE REVIEW ........................................... 12

Principle of Least Effort .......................................................... 12

SERP Studies ....................................................................... 15

Document Triage ................................................................... 16

Abstracts and Meta Descriptions ........................................... 24

Relevancy ............................................................................. 26

Google Scholar ..................................................................... 28
Times Per Participant

Appendix H. AOI Meta Description ................................................................. 151

Times Per Participant

Appendix I. AOI Cited Times Per Participant .................................................. 152

Appendix J. Number of Surrogate Document Transitions ................................. 153

Appendix K. Total Time on Task Per Participant .............................................. 154

Appendix L. Relevancy Accuracy Per Participant ............................................. 155

Appendix M. Survey Responses ..................................................................... 156

Vita ............................................................................................................. 160
### TABLE OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Anatomy of Google’s SERP</td>
<td>4</td>
</tr>
<tr>
<td>2.1</td>
<td>Loizides Funnel Model</td>
<td>20</td>
</tr>
<tr>
<td>2.2</td>
<td>Google Scholar SERP Design</td>
<td>31</td>
</tr>
<tr>
<td>3.1</td>
<td>Control Google Scholar Result Example</td>
<td>38</td>
</tr>
<tr>
<td>3.2</td>
<td>Experimental Google Scholar Result Example</td>
<td>39</td>
</tr>
<tr>
<td>3.3</td>
<td>Areas of Interest</td>
<td>41</td>
</tr>
<tr>
<td>3.4</td>
<td>Coffee Control and Coffee Experimental Relevancy Scores</td>
<td>53</td>
</tr>
<tr>
<td>3.5</td>
<td>Tea Control and Tea Experimental Relevancy Scores</td>
<td>53</td>
</tr>
<tr>
<td>4.1</td>
<td>Heat Maps for the Control Display for Participant Groups 1-4</td>
<td>62</td>
</tr>
<tr>
<td>4.2</td>
<td>Heat Maps for the Experimental Display for Participant Groups 1-4</td>
<td>63</td>
</tr>
<tr>
<td>5.1</td>
<td>Mean Time per AOI Visit</td>
<td>79</td>
</tr>
<tr>
<td>5.2</td>
<td>Group 1 Control and Experimental Heat Maps</td>
<td>81</td>
</tr>
<tr>
<td>TABLES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------</td>
<td>------------------</td>
<td></td>
</tr>
<tr>
<td>Table 3.1 Counterbalancing</td>
<td>.......................................................... 40</td>
<td></td>
</tr>
<tr>
<td>Table 3.2. Research questions and data analysis</td>
<td>.......................................................... 45</td>
<td></td>
</tr>
<tr>
<td>Table 3.3 Counterbalancing</td>
<td>.......................................................... 50</td>
<td></td>
</tr>
<tr>
<td>Table 4.1 Question 1: How old are you?</td>
<td>.......................................................... 58</td>
<td></td>
</tr>
<tr>
<td>Table 4.2. Question 3: Approximately (best estimate) how many credit hours have you completed, (not including this semester)?</td>
<td>.......................................................... 59</td>
<td></td>
</tr>
<tr>
<td>Table 4.3 Question 4: Approximately (best estimate) how many credit hours have you completed, (not including this semester)?</td>
<td>.......................................................... 59</td>
<td></td>
</tr>
<tr>
<td>Table 4.4 Question 6: How familiar are you with Google Scholar?</td>
<td>.......................................................... 60</td>
<td></td>
</tr>
<tr>
<td>Table 4.5 Total time on task: Correct perceptions vs. total perceptions?</td>
<td>.......................................................... 70</td>
<td></td>
</tr>
<tr>
<td>Table 4.6 Confidence: Correct perceptions vs. total perceptions</td>
<td>.......................................................... 71</td>
<td></td>
</tr>
<tr>
<td>Table 4.7 Document Transitions: Correct vs. total perceptions</td>
<td>.......................................................... 72</td>
<td></td>
</tr>
</tbody>
</table>
Abstract

By focusing on the point where the document triage process interacts with a search engine results page (SERP), this experiment extends our knowledge about both SERP design and document triage behavior. Prior SERP work has shown that longer meta descriptions in SERPs improve people’s ability to answer information based questions, while document triage research has shown the importance of abstracts in making relevancy decisions. Using eye tracking equipment this work employed a repeated measure within factors experimental design method replacing the existing Google Scholar (GS) SERP meta descriptions with the abstracts of the corresponding retrieved articles. Undergraduate freshmen participants were asked to use two different GS SERPs, one with a control design and one with the experimental design and determine which resources are relevant to their assigned research task.

The findings show that the participants changed how long they looked at the expanded meta description, while noticeably reducing how long they gazed at other parts of the page supporting other research findings. The addition of abstracts changed user behavior by reducing how often they made surrogate level document transitions, but did not change how often they sought out full-text documents, supporting the principle of least effort. The addition of abstracts did not contribute to changes in total time on task or participant’s relevancy accuracy. This study’s findings conflict with other work that found that longer meta descriptions corresponded with a reduction in total task time and an improvement in accuracy for informational tasks. Further research is needed to determine if this conflict was due to task differences or if the document triage task was not challenging enough.
CHAPTER 1 - INTRODUCTION

This research sought to achieve a better understanding of how undergraduate freshmen experience the document triage process as it relates to search engine result pages (SERPS) in order to better inform educators, SERP designers, and educational institutions how design affects human information behavior. The findings presented through this study show how the addition of an abstract to an academic SERP changes user behavior, supports prior document triage and SERP research, as well as uncovering topics for future study.

Statement of the Problem

Information is a constant need in scholarly research, but when there is so much information within reach, how does one find and make sense of the resources that are needed without drowning in the vast number of options available? Since the advent of the printing press and on through to the age of the Internet, new technologies have expanded the amount of information scholars potentially have available to them. Each new technology brought with it added complexity, new detachment, and the fear of excess (Gleick, 2011, p. 398). Finding relevant information can be a particularly daunting task as the number of articles published every year has been steadily increasing at an average rate of 3.45% per year from 1800 until 2003 (Mabe, 2003). This 3.45% growth rate equates to 1.4 million peer reviewed articles being published in 2009, which breaks down to one being published every 22 seconds in that year alone (Campbell & Meadows, 2011). In 1998 there was a call across all disciplines for rethinking the number of journal articles being published, seeking quality over quantity, due to the increase in articles published being nearly tripled in that decade alone (Zemsky, 1998). Today there is no evidence that the rate of peer reviewed publishing is decreasing.

Undergraduate freshmen struggle with engaging research material as they transition from
high school to college level research and writing. Head and Eisenberg (2010) found that for over three-fourths of surveyed undergraduate students, the most difficult element of studying was course-related research (p. 3). Of these students 61% claimed filtering through irrelevant results was a challenge. Another study focusing on first-term freshmen found they faced challenges searching through research information systems as well as evaluating, integrating, and applying the sources they researched. Over half (57%) of these freshmen felt overwhelmed by the vast amount of irrelevant results their online searches usually returned (Head, 2013). Yet another study from the same year showed similar results, finding freshmen had trouble making sense of the information they found using library search tools (VanderPol, Swanson, & Kelly, 2013). Undergraduate freshman are struggling to make sense of the results they are finding in relation to their course work.

As Gleick points out, our best hope against the flood of information and its challenges can be summarized in two words: filter and search (Gleick, 2011, p. 409). Typically searching for and filtering through academic articles is commonly done with the aid of an information retrieval (IR) tool, but as the research discussed above found students struggle with filtering and making sense of what their searches return. IR tools first came into being shortly after the Second World War to address the information explosion that scientific researchers were facing at that time (Saracevic, 2008). The earliest conception of an IR tool is accredited to Vannevar Bush, who conceived of a Memex machine that would store vast amounts of records that could be consulted quickly by researchers (Bush, 1945, p 2). IR systems emerged as static batch systems until the 1970s, when the combination of computer systems and communication systems made IR tools dynamic and interactive (Spink & Saracevic, 1997, p. 741). Web-based search engines and digital databases of academic journals like those provided by EBSCO and other vendors are
examples of contemporary IR systems or tools. Over the years, along with technology improvements, testing and experimenting with the algorithms that drive the document search and retrieval within these systems has been improved on the system’s side (Saracevic, 2008, p. 780). But while IR tools and search engines are improving, undergraduate freshmen are continuing to struggle to make sense of the results these tools retrieve.

Research that focuses on the interaction between humans and computers is known as the field of human computer interaction (HCI). There are two areas of HCI research that both relate to the information retrieval tool design and the human interaction with the IR tool and the results it retrieves: Search Engine Results Page (SERP) and an area of information seeking behavior known as document triage. SERP studies investigate the design of what users see on the screen when using a search engine and how design affects their behavior and understanding of the resources they have discovered. Document triage investigates what behaviors are common while people make sense of what is and is not relevant to their information needs. Considering the findings showing the frustration and challenges faced by undergraduate freshman, this area of search engine display design was investigated in search of improvements.

**Statement of Purpose**

Considering the sense-making problem confronting undergraduate freshmen, this research investigated the behavioral effects a SERP meta description design change would have on the undergraduate document triage process. By focusing on this point where the document triage process interacts with a SERP, this proposed study intended to extend our knowledge about both SERP design and document triage behavior. The main objective of this research was to investigate the effects of a yet unstudied SERP design change in Google Scholar (GS), a search engine that relates to academic articles. Due to the work showing positive results from
lengthening the meta description (Cutrell, E., & Guan 2007) and the studies showing the importance of abstracts (Loizides, & Buchanan, 2011; Saracevic, 1969, Cool et al. 1993), this study replaced the existing GS meta descriptions with the abstracts of the retrieved articles.

SERP Meta Descriptions and Document Triage

Anyone who has used an online search engine, like Google or Yahoo!, to surf the Web or an online database like those found through library webpages, has seen a SERP. A SERP usually contains about ten results and displays the title of each Web page or resource retrieved by the search engine along with other possibilities, like each of the retrieved resources web page’s uniform resource locator (URL), and short descriptions of what the results page is about or a snippet of what it contains, commonly called a meta description (See figure 1.1). Some SERPs also have other elements, like search options and advertisements.

![Figure 1.1: Anatomy of Google’s SERP](image)

The meta description of a SERP contains information that gives the reader information about a Web page’s content. One study showed that adding longer meta descriptions to SERPs improved the study’s participants ability to accurately answer questions using search engines (Cutrell & Guan, 2007). This finding indicates that altering meta descriptions in scholarly SERPs
by lengthening them could be a help to freshmen undergraduate students. But academic articles are different from webpages, which is why findings from document triage research are useful in informing what sort of information in the SERPs meta description could be helpful to freshmen researchers.

Bae et al. (2005) defined 'document triage' as “the practice of quickly determining the merit and disposition of relevant documents – including web pages, periodical articles, and other published materials – that one may locate using a search engine, receive from an automated delivery mechanism, or obtain from a human intermediary” (p. 130). It is this definition of document triage that will be used for this study. Document triage studies focus more on the human information behavior side of how people make sense of what is relevant when evaluating document sets, but also study aspects of interface design (Saracevic, T., 1969; Bishop 1998; Badi et al. 2006; Bae et al., 2006; Bae et al., 2005; Bae et al., 2010; Buchanan & Loizides 2007; Buchanan, & Owen, 2008; Loizides & Buchanan, 2009; Loizides & Buchanan, 2011, Loizides, 2012). The document triage findings most related to this study have to do with an overall model for the document triage behavior, the Loizides ‘funnel model’ (Loizides, 2012), which describes the document triage process for people using search engines. This model breaks document triage with IR tools into three levels, the ‘surrogate triage stage’, the ‘within document triage stage’, and the ‘further reading triage stage’ (Loizides, 2012). These terms and stages will be discussed at length in the literature review. This study is mainly concerned with the first of these three stages which involves SERPs. Several of these document triage studies show that abstracts of academic articles play a large role in the decision making process, in that they are one of the areas of an academic paper that receive the most attention while people make relevancy decisions (Loizides, & Buchanan, 2011; Saracevic, 1969, Cool et al. 1993). Relevancy itself is a
rich term in the field of information science, but for the purposes of this study it will be defined as "pertaining to the matter at hand." According to Saracevic (1996) this is not only the meaning of relevance defined in major dictionaries, but also the meaning intuitively understood by people everywhere, effortlessly applied without any need for a more formal understanding of what 'relevance' is. Because of these document triage findings and their relationship to SERP research, the GS SERPs’ meta descriptions were replaced with full abstracts in this study.

**Significance of the Research**

While many aspects of GS, like the resources it indexes (Arlitsch, & O'Brian, 2012), its perceived usefulness to end users (Cothran, 2011) and popularity (Herrera, 2011) have been researched, GS’s SERP has yet to be studied in this way. While other studies have investigated SERP areas of interest (Dumais, Buscher, & Cutrell, 2010; Joachims et al., 2005; Lorgio et al., 2008) these studies have not investigated gaze attention on display elements unique to GS, like author and publisher information and the number of citations and links to other works. Nor have there been studies on how these areas of interest could possibly affect document relevancy perception. This study investigated these areas of interest, as well as article titles and meta descriptions, informing us of their gaze effects on an academically orientated SERP.

Considering the research about the challenges faced by freshmen undergraduate students with filtering through and making sense of research results (Head & Eisenberg, 2010; Head, 2013; Van der Pol, Swanson, & Kelly; 2013), this population would be the one selected for study to discover if the design change is beneficial. This studies’ findings lead to more insight into how SERP design can improve document triage to help alleviate the problems that undergraduates face or encounter in making sense of the ever-growing amounts of information they find during their research process.
Experimental Design and Rational

This study employed a repeated measure within factors experimental design. The independent variable was changed in the meta-description within GS’s SERP. Undergraduate participants were asked to view two different SERPs and navigate them as they would normally to determine which resources were relevant to their assigned research task. One SERP was in the control format and the other in the experimental format. Participants were asked to examine documents shown on the SERP that they rated for relevancy in both the original and the new design environments while their session was being recorded with eye tracking software. The eye tracker is an unobtrusive tool that was used to track user visual attention, navigation patterns, and time on task. The eye tracker also allowed actual users to be engaged in research tasks and to triage as they would normally. While the experimental design placed some limitations on user behavior and focus on a limited step of the overall information-seeking process, the experimental method is welcomed in HCI because it allows for hypotheses to be tested (Reader & Payne, 2007). The dependent variable data sources collected were the mean length of time spent visual in the eye tracking areas of interest, mean triage stage transitions, total time on task, and relevancy scores. The repeated measures within factors design also allowed for feedback from participants to be collected about their experiences within each environment in the form of a post-experiment survey.

Research Questions

In order to investigate the effects adding articles’ abstracts to GS’s SERP design, the following research questions were constructed to guide the experiment. The central research questions for this study are:

RQ1: How does augmenting the meta description with an abstract in a SERP affect document
triage behavior in undergraduates?

RQ2: How does augmenting the meta description with an abstract in a SERP affect document triage behavior in undergraduates at the surrogate triage stage?

RQ3: How does augmenting the meta description with abstracts in a SERP affect perception of document relevancy?

RQ4: How does augmenting the meta description with abstracts in a SERP affect document triage efficiency?

These questions were answered through three primary data sources collected during the experiment: the data collected by the eye tracker, the participant relevancy responses, and the post experiment survey. Together all of these data will help answer the study’s first and overarching research question: how does replacing the meta description with an abstract in a SERP affect document triage behavior in undergraduates? Specifically the experiment recorded eye tracking gaze plots, possible differences in mean participant time looking at areas of interest, and the mean number of links to other documents. These data, along with questions from the post-experiment survey, were used in answering the second research question: how does augmenting the meta description with an abstract in a SERP affect document triage behavior in undergraduates? The accuracy of the participant’s relevancy sorting compared to that of a subject matter expert in each results set along with questions from the post experiment survey will answer the third research question: how does replacing the meta description with an abstract in a SERP affect perception of document relevancy? The fourth and final research question, how does augmenting the meta description with abstracts in a SERP affect document triage efficiency, was answered by recording total time on task in each set of documents along with combined analysis of the accuracy of the participant’s relevancy, answers from the survey, and
the number of links to other documents. For the purposes of this study efficiency will be defined as the ratio of time to accuracy of the relevancy-sorting task.

**Expected Outcomes**

Based on prior research findings it was hypothesized that the design change will produce results similar to the study (Cutrell & Guan, 2007) that augmented the meta description of non-academic search engine SERPs with a longer meta description. It was hypothesized that adding abstracts to GS’s SERP would increase both gaze time in the meta description areas and improve participant ability to sort documents more accurately into groups of similar relevancy when compared to that of a subject matter expert.

It was also hypothesized that adding abstracts to a SERP, the surrogate document triage level in Loizides’ (2012) funnel model, would change a user’s information seeking behavior and perhaps relevancy choices. Having abstracts at the first level of the document triage process rather than having to navigate to another page at the surrogate level or to the second within document triage stage via hyperlinks and back buttons could reduce how often participants leave the SERP while making relevancy decisions reducing the amount of navigation to other surrogate documents, like the webpages GS links to, and links to Full Text documents available on the SERP. It was also hypothesized that adding the abstracts into the search results would not change total time participants spend to complete the document triage task, as reading and skimming times should be similar and the times savings in load times, while perceivable, are not likely to be significant.

**Conclusion**

The number of academic resources available for research is constantly growing at a rapid rate. Freshmen undergraduates have been identified as a group who struggle with making sense
of search results when doing research for class assignments. Prior work has shown that longer meta descriptions in SERPs improve people’s ability to answer information based questions (Cutrell & Guan, 2007). Other research has shown the importance of abstracts (Loizides, & Buchanan, 2011; Saracevic, 1969, Cool et al. 1993) in making relevancy decisions. This study replaced the existing GS SERP meta descriptions with the abstracts of the corresponding retrieved articles and tested the effects of the SEPR design change with an experiment. The study brings new understanding of how the design change affected gaze time in areas of interest, such as the title, author and publisher information, the meta description, and the number of citations and links to other works. This study also brings new insight to how this design change affected undergraduates perception of which SERP results were relevant and how it changed their document triage behavior.

Definitions

Document Triage - The practice of quickly determining the merit and disposition of relevant documents – including web pages, periodical articles, and other published materials – that one may locate using a search engine, receive from an automated delivery mechanism, or obtain from a human intermediary.” (Bae et al., 2005, p. 130).

Meta Description - The automated descriptions found on SERP after displayed results titles, generated by search engines that take into account both the content of a page as well as references to it that appear on the web. The goal of the meta description is to best represent and describe each result and explain how it relates to the user's query.
Principle of Least Effort - According to Zipf (1949), people naturally adopt a course of action that keeps their usage of resources in harmony, that is, managing their behavior in such a way that they will expend the least average amount of effort as the accomplish their tasks.

Relevancy - Saracvevic (1996) points out that there is one meaning that is understood best by most people and this definition states that relevancy means anything “pertaining to the matter at hand” (p.3).

Satisficing - The tendency for humans to seek a solution with an acceptable quality of completeness, rather than to strive to obtain an ideal or perfect outcome. (Case, 2012, p. 389).

SERP - An acronym standing for search engine results page, the page of results generated by a search engine like Google or Summon.
CHAPTER 2 – LITERATURE REVIEW

As outlined in the introduction, this study aimed to investigate a particular problem relating to information seeking behavior, namely how does a design element of a search engine display affect an undergraduate’s ability to be an effective researcher? In order to investigate the design change to Google Scholar’s meta description, two frameworks have been identified. The first is search engine results page (SERP) studies. SERP studies are focused on web-based IR tools and their design. The second is document triage, a framework relating to human information behavior that directly involves information retrieval (IR) tools. These two frameworks draw on literature taken from three different areas of the information science field, HCI, human information behavior (HIB), and information storage and retrieval (ISR). This review of the literature will show how these areas of study support each other and provide useful experimental methods that informed this study’s design. We begin with a review of the literature on the Principle of Least Effort, a key HIB theory that explains the document triage behavior, an ISR framework, and follow with summaries of SERP and document triage studies, which are also considered to be HCI studies. These sections will be followed by other important topics, like a discussion of abstracts, relevancy, and prior research on GS. The review of literature ends with discussion of prior SERP and document triage methodologies and findings that inform this project and its methodological design.

Principle of Least Effort

The term Principle of Least Effort was first coined by a French philosopher named Guillaume Ferrero. Ferrero (1894) theorized that without external influence humans would remain in an inert mental and physical state. It was however, philologist George Zipf (1949) who became more famously associated with the theory. Zipf found that individuals take the course of
action that will involve the probable least average of their work, or in other words the course of action that would take the least effort. Zipf’s work investigated the statistical distributions of words in works of literature, and the principle is now referred to as Zipf’s law, in the context of the frequency of word distribution in human languages (Manning, 1999). This principle of the exertion of least effort has been found in other fields as well.

Psychologists have also investigated this principle and found evidence to support it. There has been a long-standing behavioral principle that both people and animals avoid physical effort when all else is equal (Hull, 1943). The principle for least effort has not been directly investigated for cognitive effort, although cost-benefit analysis studies (Kahneman & Tversky, 1979; Stephens & Krebs, 1986) have shown support for the Principle of Least Effort with indirect evidence, but the costs evaluated were beyond the confines of cognitive demand alone. Kool et al. (2010) conducted a series of experiments to investigate the effects of anticipated cognitive demands and to provide direct empirical evidence for or against this principle. They found support for the idea that anticipated cognitive demands play a significant role in behavioral decision-making, providing more evidence to support the Principle of Least effort in the cognitive domain as well as physical effort. These findings suggest that, psychologically, people avoid not only physical effort, but mental effort as well.

In the field of economics, this principle is better known by the term *satisficing*. This term is generally used now to mean a behavior where a selection is made based on something being *good enough* or the minimum satisfactory condition or outcome rather than an optimal solution. The word *satisfice* was created in 1947 by Herbert A. Simon. Later Simon (1956) refined his definition giving the word its current meaning. Simon derived the word from his studies of economics and the way managers make decisions, he claimed that, “administrative theory is
peculiarly the theory of intended and bounded rationality—of the behavior of human beings who *satisfice* because they have not the wits to *maximize*” (Simon, 1997, p. 118).

Satisficing has been used in information science studies by a variety of researchers including Warwick, Rimmer, Blandford, Gow, and Buchanan (2009) and Connaway, Dickey, and Radford (2011) to explain why users did not spend much effort and would settle on resources rather than looking further in information-seeking behavior. Recently Duggan and Payne (2009, 2011) and Reader and Payne (2007) have shown a satisficing account of the skimming process that allows readers to focus on the most important elements of a single document. Wilkinson, Reader, and Payne (2012) also supported these findings, showing that satisficing to make sense of texts is a widely used strategy in skimming, though not the only one.

Information scientists have made findings that support the Principle of Least Effort and consider it a more general theory for the behavior of information seekers. Wilson (1977) made the observation that individuals look for “a new combination of sources (which might be all of the old sources plus one additional new source) that will at the least cost, bring our information and advisory systems up to a level of adequacy” (p. 73), when investigating adaptive changes in information gathering habits. Poole (1985) analyzed information seeking literature and found that 40 of the 51 studies he sampled lent their support to the Principle of Least Effort. Dervin (1983) noted in her work that the reliance of people on close friends and relatives as information sources was a demonstration of a “law of least effort” (p. 158). Mann’s (1993) work also investigated the ways information seekers exploit systems to minimize effort. Recent investigation into the interpersonal means of seeking information (Xu et al., 2006) and the use of library research interface (Connaway, Dickey, & Radford, 2011) continue to support this Principle of Least Effort. This principle of finding strategies to minimize effort is the main
theory that drives document triage behavior, which often involves the information seeker making use of search engine IR tools.

**SERP Studies**

SERP research has investigated features of search engine results pages using eye-tracking equipment (Dumais, Buscher, & Cutrell, 2010; Lorgio et al., 2008; Cutrell & Guan, 2007; Pan et al., 2007; Joachems et al., 2005). Studies showed that users view SERPs roughly in a linear top down manner (Cutrell & Guan, 2007; Joachems et al., 2005). Dumais, Buscher, and Cutrell (2010) investigated the effect that other parts of SERP like advertisements, search boxes, and other elements of Google’s SERP had on gaze attention. They found participants fell into three groups of search behaviors, those whole look at a SERP broadly, others who focus only on the results, and the third who regularly look at advertisements.

Many SERP studies have been concerned with effect that the rank ordering of results have on users and their perceptions of relevancy. These studies found that information seekers have strong biases toward results ranked at the top of the results list (Lorgio et al. 2008; Pan et al., 2007; Joachems et al., 2005). This bias is so strong that undergraduates trusted the rank order over seeing that the meta description, the brief synopsis under the result title containing bolded key words, did not match with their information seeking tasks (Pan et al, 2007). Another rank order SERP study showed that the change in rank order increased scrutiny of the meta descriptions shown with the results, as well as finding that 96% of students only viewed the top ten results (Joachims et al., 2005). The first three results receive the most attention out of ten, with the top two receiving the most attention overall (Lorgio et al. 2008; Pan et al., 2007).

While most studies only investigated results as they are normally displayed, a few studies changed the rank order of results only (Dumais, Buscher, & Cutrell, 2010; Lorgio et al., 2008;
Pan et al., 2007; Joachims et al., 2005). One study by Cutrell and Guan (2007) altered the length of the meta description to see if it would affect information seeking tasks. Their results showed that the longer the meta description the more time subjects looked at those areas. Longer meta descriptions improved the accuracy of information tasks, tasks that asked participants to answer questions with search engines, but lowered performance with web navigation tasks, those that asked the participants to find particular Web pages (Cutrell & Guan, 2007).

The cognitive processes that steer the behavior involved with the information and navigation tasks have not yet been explored in SERP research (Lorgio et al., 2008). Also SERP studies mainly focused on Google and Yahoo! search engines (Dumais, Buscher, & Cutrell, 2010; Lorgio et al. 2008; Cutrell, E., & Guan 2007; Pan et al., 2007; Joachems et al., 2005) and are not generalizable (Pan et al., 2007). Beyond experimenting with rank order and meta descriptions for web page searches, no research has been done with design changes of other elements like URL location, or citation and publisher information elements like those found in GS. This study is the first to investigate the effects other types of SERP features that are more related to academic results rather than those relating to the full World Wide Web. It is also the first study to experiment with GS and longer meta descriptions, but also academic article abstracts as meta descriptions, a design improvement taken from the findings of document triage research.

**Document Triage**

The term 'document triage' is a relatively new addition to the information science field and has been previously used in a limited number of works but continues to grow in popularity, recently being used in the study of medical databases (Wei, Kao, & Lu, 2013) and the field of forensics (Bogen et al., 2013). The term 'document triage' has undergone some evolution and
was first introduced as ‘information triage’ (Marshall & Shipman, 1997). In a later work Bae et al. (2005) defined 'document triage' as “the practice of quickly determining the merit and disposition of relevant documents – including web pages, periodical articles, and other published materials – that one may locate using a search engine, receive from an automated delivery mechanism, or obtain from a human intermediary” (p. 130). Later, Loizides and Buchanan (2009) defined document triage as: ‘the moment in the information seeking process when the user first decides the relevance of a document to their information need” (p. 139). Fernando Loizides built two models for the behavior in his 2012 dissertation. Here Loizides (2012) uses a slightly different definition for document triage; “Document Triage is the fast process by which information seekers go through a set of potentially relevant documents to establish relevance to their information need’ (p. 45). Because this study investigated GS’s SERP and how a design change affects document triage, for the purposes of this study we will use the Bae et al (2005) definition which is more precise rather than Loizides’, which is less clear as to what is meant by a document set.

Document Triage Models

While commonly classified as an ISR studies in research databases document triage looks to information seeking models to define it theoretical bases. Several information seeking models were explored in relation to document triage and information retrieval tools to support the creation of the two models (Loizides, 2012). Each information-seeking model that was used was selected for strengths in representing different qualities of the information seeking process. Ingwersen’s (1996) integrative model was used to frame the overall cognate process involved, but Saracevic (1996) and Wilson (1999) pointed out the model’s limitations in testability. Loizides also used Bates’ (1989) model to acknowledge the difference in active and passive
information seeking models. Loizides draws most from Ellis’ (1989) model to show links between document triage behavior and other information seeking behavioral aspects. Marchionini’s (1997) model is used to inform document triage’s information seeking procedures. Loizides uses Kuhlthau’s (1991) model to cover cognitive and emotional elements of information seeking. Within these models Loizides identifies the stages with these models that relate to document triage and where he places his document triage models. The crucial component of these three models to document triage is relevance decisions. In Ellis’s model this is the ‘differentiating’ step, in Marchionini’s model this is the ‘examine and extract results’ step, and in Kuhlthau’s model this is the ‘formulation’ step.

The first model Loizides (2012) creates is a high level model, which he names the ‘funnel model’, which conceptualizes the document triage information seeking sub-process as a whole (see Figure 2.1). The model contains three levels of triage that information seekers undergo during document triage; in-depth reading cannot be considered a part of the document triage process. In Loizides’ (2012) ‘funnel model’ the document triage process begins as soon as a set of documents is presented to an information seeker, usually by an information retrieval tool like a search engine. The triage process ends after the information seeker has evaluated all these documents, and their relevance to the information seeker’s need has been ascertained (Loizides, 2012, p. 172). Even if documents in a results list are never reached, they are still considered to have been implicitly rejected with regards to that specific triage occurrence. There are three stages at which a document can be evaluated: the surrogate level, the within-document level, and the further reading stage (Loizides, 2012).

**Document Triage Findings**
Bae et al. (2006) found common strategies for the overall process of performing document triage. The first commonly identified strategy was that many participants focused on reading early in the task, and held the organizing portion of the task until later in the triage process. Other participants used a strategy where they organized in tandem with skimming. The final strategy used by the last category of participants organized articles only using document surrogates, such as the results display in a SERP, rather than spending time reading.

Many studies (Loizides, & Buchanan, 2011; Saracevic, 1969, Cool et al. 1993) pertaining to document triage have investigated only the triage of academic papers, which are structured with titles, abstracts, and section headings. These specific document areas such as title, abstract, and section heading text, have been shown to receive the most attention (Loizides, & Buchanan, 2011; Saracevic, 1969, Cool et al. 1993) during the document triage process. These same aforementioned studies also showed that these elements receive a disproportionate amount of attention compared to the space they occupy in a document, implying that these are the portions of the document most used to discern its relevancy. 

When investigating document triage behavior of PDF and printed documents Loizides and Buchanan (2009, 2011) found three patterns participants typically used when navigating a particular document during the triage process. Most use a ‘step up navigation’ approach, which is when the participant linearly skims the document from top to bottom, stopping at areas of interest. About 10% used a hybrid behavior, switching between the ‘flatline navigation’ approach (when a participant only looks at the first page) or ‘begin and end approach’ (when the participant looks at the first page and then jumps to the conclusions only) (Loizides & Buchanan 2009). Morti and Shinoda (1994) and Bae et al. (2005) found that time spent in a document positively affected relevance ratings. This same study by Bae et al. (2005) also found that the
higher the number of scroll events to a particular document was an indicator of a higher perceived document value.

Figure 2.1: Loizides Funnel Model

Loizides’ (2012) ‘surrogate triage’ stage is when the information seeker is likely to have first contact with documents surrogates, such as a results list supplied by an information retrieval tool. Document surrogates most commonly refer to a representation of a document such as the title and a short summary, like the elements of a results list like those supplied by an IR tool, be it a search engine, database, federated search engine or an online public access catalogue (OPAC). An academic journal’s web page relating to an article would be considered a document surrogate if it only contained an abstract or summary. If the web page contained the full article then it would be considered a document for the purposes of this study (Loizides & Buchanan, 2007). Often, this stage occurs when information seekers perform a query based search on a
search engine. Most of the relevancy decisions about the documents are made at this stage, with the majority of documents being rejected without any viewing due to the limited depth of triage a user is willing to undertake when looking through a search results list (Spink et al., 2002). This is also supported by a study showing that users are likely to make a relevance judgment on a document without opening the full text (Buchanan & Loizides, 2007). This same study also showed that most users have a linear approach to reviewing results at the surrogate level (Buchanan & Loizides, 2007).

Loizides calls the second stage of the “funnel model” the ‘within-document triage stage’ (p. 173). The ‘within-document triage stage’ begins when the information seeker chooses to view the full text of an individual document and evaluate its relevancy. During this stage, the information seeker is likely to skim some parts of the document at a fast pace. This process is multifaceted and is affected by factors like the complexity of the information need, the obscurity and layout of a document and the information seeker's triage skill (Loizides, 2012, p. 173-174). This process has enough distinct elements that Loizides’ other document triage model is concerned only with modeling the document triage behavior at this level.

Loizides’ (2012) finds that after an information seeker finishes the first two levels of triage (Surrogate Triage and Within-Document Triage) they are left with the set of documents that they have found to be relevant to their information needs. Due to how brief these relevancy decisions have been made, one study showed that the relevance decision is made in less than one minute (Buchanan & Loizides, 2007); the true relevance of the remaining documents will have to be made when the documents are actually read. This is the stage, called the ‘further reading stage,’ when the documents may actually be read rather than skimmed or parsed through for bits of information. According to Loizides, (2012, p. 174) this stage has not been studied to date.
This study will be investigating how a change at the surrogate level of an information retrieval tool’s display will affect an information seeker’s behavior during the first two stages of Loizides’ ‘funnel’ document triage model.

As Loizides’ (2012) ‘funnel model’ suggests, triaging electronic documents is a multifaceted process (Badi et al., 2006) and involves both the use of document surrogates, like the results listings found in SERP, and documents in web page HTML or PDF format. Researchers within the literature of document triage found that active reading while searching through documents is usually brief (Adler et al., 1998). Loizides and Buchanan (2011) found that when time was limited to determine relevancy, there was no difference in the significance rating of documents under time pressure compared to those triaged with no time limits to determine relevancy. It has also been shown that there are significant error rates when executing the document triage tasks (Buchanan & Loizides, 2007).

**Document Triage and SERP**

Bae et al. (2006) found common strategies for the overall process of performing document triage. The first commonly identified strategy was that many participants focused on reading early in the task, and held the organizing portion of the task until later in the triage process. Other participants used a strategy where they organized in tandem with skimming. The final strategy used by the last category of participants organized articles only using document surrogates, such as the results display in a SERP, rather than spending time reading.

Many studies (Loizides, & Buchanan, 2011; Saracevic, 1969, Cool et al. 1993) pertaining to document triage have investigated only the triage of academic papers, which are structured with titles, abstracts, and section headings. These specific document areas such as title, abstract, and section heading text, have been shown to receive the most attention (Loizides, & Buchanan, 2007).
2011; Saracevic, 1969, Cool et al. 1993) during the document triage process. These same aforementioned studies also showed that these elements receive a disproportionate amount of attention compared to the space they occupy in a document, implying that these are the portions of the document most used to discern its relevancy.

So far there are still many unanswered questions about the document triage behavior. Much of the document triage work has been centered on academic articles (Loizides & Buchanan 2009, 2011). Considering the structure of these articles and their use of headings, within-document triage patterns could be different for other types of documents like news articles, web pages, and books or ebooks. There also is little document triage research about how users choose between available books in online public access catalogs, like those provided by libraries, (Hinze et al. 2012) or when searching for non-text based documents like images.

The cognitive reasons driving the behavior has not yet been researched in document triage studies (Loizides, 2012). Also no known work has been done studying the effect that experience and training have on the document triage process. The differences in display design has only begun to be studied, with most document triage studies focusing on PDFs or software not readily available to the public like the Virtual Knowledge Builder (Badi et al., 2006; Bae et al., 2006; Bae et al., 2005), Redeye Analysis Workbench (Bogen et al., 2013) and Pubtator (Wei, Kao & Lu, 2013). While the document triage is a useful perspective on the behavioral process of a select portion of the information seeking process, it is not all-encompassing nor a widely known perspective within the human information behavior field.

Despite document triage being a new term and the behavior models a new perspective on a limited aspect of the information seeking process, these findings and perspectives hold value and practical promise in further understanding how people search for information. It is because
of this focused perspective on this sub-process of the overall information seeking process that gains can be made by adapting our technologies to what we are now beginning to understand about existing behavior within current information seeking environments, like GS’s SERP. This study will provide insight as to how much information is useful to undergraduate researchers at the surrogate level of the triage process.

Abstracts and Meta Description

Abstracts are information packets that, when well-written, convey a document’s subject content (Cleveland & Cleveland 2001, p. 101-102). The purpose of this conveyance of subject matter is to allow people to be able to understand what a document is about from reading the abstract alone. This article ‘aboutness’ that an abstract is concerned with has to do with what words say and what they mean; unfortunately computers are not good at understanding aboutness (2001, p. 98). While a computer can find keywords and retrieve documents, people still need to read through and decide what the retrieved documents are actually about. A good abstract should convey what an article is about and its key findings.

Abstracts have different types: indicative abstracts, informative abstracts, indicative-informative, and critical abstracts. The indicative abstract describes the type of article it is and what it is about without telling readers the contents in detail (Zhang & Liu 2011). The informative abstract reports the details of the research, usually containing as much qualitative and quantitative data as it can, bridging ideas and data (Cleveland & Cleveland 2001, p. 57). Indicative-informative abstracts are a type of hybrid that present data details and cover other aspects in the form of descriptive statements (Zhang & Liu 2011). Critical abstracts are more like reviews or editorials of an article and have debatable usefulness due to the greater possibility of bias, but maintain usefulness if well written by a knowledgeable author (Cleveland & Cleveland
It is even possible to classify abstracts by the abstract’s author, which can be the document’s author, a subject matter expert, or a professional abstractor; of the three professional abstractors are considered to be the best at producing abstracts (Cleveland & Cleveland 2001, p. 59).

Abstracts are now considered an expected part of an academic paper, but this was not always the case. While some journals included abstracts as a part of their publications as early as the 1920s, they became more and more of a norm over time, with major journals like Nature not introducing them till as late as 1970 (Huckin, 1993). An abstract’s function has proven useful. Winker (1999), in review of medical journal abstracts, claimed the abstracts, aside from titles, were the most read portion of articles. Since the abstract’s consistent inclusion in most academic articles, the abstract itself has become a subject of study, particularly the effectiveness of structured abstracts which tend to be more efficient, easier to read, and take more space, but are no more or less accurate than unstructured abstracts (Zhang & Liu 2011). Hartley and Sydes (1997) could not find studies about the effectiveness of earlier unstructured abstracts, aside from studies about the effects summaries and overviews have on recall, but found that structured abstracts were easier to read, but bad writing could reduce that advantages of a structured abstract (p.125).

Because of the document triage findings (Loizides, & Buchanan, 2011; Saracevic, 1969, Cool et al. 1993) and other studies relating to scholarly abstracts (Zhang & Liu 2011; Winker, 1999; Hartley & Sydes, 1997), this important element of articles has been selected to be the element added to the SERP design. The abstract of an academic article makes the most sensible augmentation to a SERP’s meta description, considering that a meta description already contains a short snippet designed to give information seekers more clues as to what a Web page or
resource is about. Both an abstract and a meta description are designed to give researchers the same information, clues indicating if a resource is relevant to them or not.

**Relevancy**

As document triage is most concerned with the steps relating to the information seeker determining relevancy, the difficult concept of relevancy itself needs to be addressed. Information science tends to have two separate approaches to the term relevance. The first is determining the meaning of the word, and the second is concerned with technical measures of document retrieval.

The breadth and depth on the topic of commentary and empirical investigations of relevance span beyond the boundaries of the field of information science alone, covering perception, attention, memory, persuasion, attitude and belief creation. Besides, research on human psychology, mass media and advertising studies have explored the topic (Case, 2012, p. 105).

Relevancy itself seems to be an easy topic to understand but a difficult one to define. According to Saracevic (1975) “information science is not the only subject concerned with the notion of relevance. For two thousand years logicians and philosophers have been struggling with the notion (p. 144).” Furner (2004) suggests that understanding relevance is the best way to understand what information is. Ritchie (1991) offers a definition for relevance that attempts to tie the two meanings together: “Relevance has to do with meaning and describes the relationship of patterns and whatever patterns indicate to the cognitive environments of the originator and perceiver of a message. Patterns with communicative potential are data, and data with relevance are information” (p. 20). In regards to relevance and communication, Sperber and Wilson (1995) construct a broad meaning for the term and write about relevance being any patterned message
that attracts our attention and calls for interpretation. Saracevic (1996) points out that there is one meaning that is understood best by most people and this definition states that relevancy means anything “pertaining to the matter at hand” (p.3). While this may be the more intuitively understand meaning of the word, this definition is often not enough for finer applications of the term:

From intuitive understanding of relevance we can derive that it has attributes such as: it is based in cognition; it involves interaction, frequently communication; it is dynamic; it deals with appropriateness or effectiveness; and it is expressed in a context, the matter at hand. When applied in scholarly and scientific realms, used generally, it also has specific meaning in theoretical or empirical constructs derived in various fields (Saracevic, 1996, p. 203).

Because of this intuitive understanding he argued that the: “treatment of relevance in information science must follow intuitive use of relevance” (p. 203).

Within the field of information science, Saracevic (1996) argues along the same lines as Ritchie (1991), but goes further to create five relationship-based categories of relevancy. Systemic or algorithmic relevancy is concerned with the relationship between a question and information objects like documents. Topical or subject relevance is the relationship between the subject desired and the subjects retrieved by or in existence. Cognitive relevance or pertinence is the relationship of states of knowledge and user need with retrieved documents or documents in existence. Situational relevance or utility is the relationship between the problem at hand and documents retrieved or in existence. Finally, motivational or affective relevancy is concerned with how motivations of the user relate to retrieved documents or documents in existence.

Clearly, relevancy is a rich term with many definitional aspects worth considering. In information science, relevance is more closely concerned with ‘aboutness’ and information needs matching up with documents or other information resources satisfying that need. If aboutness, or topicality, meets the need then the information resource is considered to be relevant or “on the
topic” (Harter, 1992). How relevancy is determined though electronic means and the algorithms and logic associated with it is outside of the scope of this study. This study is focused on the intuitively understood human relevancy decision-making process. In other words, does a document relate to the topic at hand? Considering the proposed experiment will give participants an assigned sorting task, they will be asked if a document fits one matter at hand over another. This would best fit into two of Saracevic’s categories, situational relevance and cognitive relevance. The participant’s information needs will be contrived in order to give the participants a task, thus creating a relationship to the documents used in the study and the task given the participants. But by asking participants to sort documents relating to the task at hand they have to correctly assess if a document fits well with one category or another, which is more of a pertinence relationship. This task is actually quite similar to an assigned topic of research; where students are given a topic or topics to research, find pertinent articles, make sense of them and then write up their interpretations of their findings.

Google Scholar

Google Scholar (GS) is a freely accessible web based search engine that indexes full-text journal articles, technical reports, preprints, theses, books, and other documents, including selected Web pages that are deemed to be ‘scholarly.’ Many of GS's search results link to commercial databases and their journal articles, limiting access that many people will have to the full articles. Many of these article pages do include the article’s abstract (Jacsó, 2005). Cothran (2011) demonstrated qualitatively GS’s acceptance and use among graduate students. GS’s perceived utility is such that it is also finding a home on library web sites (Neuhaus, Neuhaus & Asher, 2008). GS is also becoming popular and is considered the preferred choice by many in the sciences and social sciences (Herrera, 2011). GS is not without its drawbacks, however. Arlitsch
& O’Brian (2012) found that Google Scholar has difficulty indexing the contents of institutional repositories, due to the fact that most repositories use the Dublin Core metadata schema, which GS cannot index properly. Only 38-48% of the surveyed documents were accessible and required using different metadata schemas in order to improve GS’s ability to access articles in the same surveyed repositories (Arlitsch & O’Brian, 2012). Another problem for Google Scholar is its access to indexes. The material GS indexes while completing its searches, like structured discipline-specific vocabularies, abstracts, and other elements, are considered proprietary information by many stand-alone database companies that provide them. These database providers are concerned that their proprietary content is becoming too available to nonsubscribers. The general idea of index-based discovery tools like GS may weaken interest in their proprietary products so much so that they may stop providing access to their indexes (Breeding 2014).

Of those surfing the Web 65% are using Google, the Webpage based search engine (comScore Inc, 2014). A 2010 survey found that 89% of undergraduates students used Web search engines when starting their research while none of those surveyed started with a library resource (Rosa et al., 2010). Because of GS’s similarity to the often-used Google, it is appropriate for use in this study. Even though it indexes different material than Google, the GS display is very similar and should not be difficult for the undergraduates to interpret. Also, web access to GS’s SERP code makes it easy to recode SERPs for the experimental testing purposes.

This study uses GS as the information retrieval tool to test this design change because it is a tool readily available to those with Internet access. GS is also a search engine that is more useful for undergraduates for class purposes than other web-based search engines because it searches for academic articles rather than other web documents. Many of the existing document
triage studies test design changes within an application called Virtual Knowledge Builder (VKB), an experimental spatial hypertext system that is not widely available or used (Bae et al., 2005; Bae, et al. 2006; Bae et al 2010; Shipman, 2004). In fact many of these studies must train participants on how to use the VKB software before running their experiment. Most SERP studies use Google or Yahoo! to test rank order changes (Dumais, Buscher & Cutrell, 2010; Lorigo et al., 2008; Guan & Cutrell 2007; Pan et al., 2007; Joachims et al., 2005) or changes in the SERP design (Cutrell, & Guan 2007). Considering this study’s interest in SERPs and finding design aids for undergraduates, it will use GS due to its availability, similar design to Google, and the likelihood students will not need prior training. Also this study will be the first experiment known to the researcher to test GS’s SERP.

GS’s SERP design is closely related to Google’s design. In 1998 when Google first launched onto the World Wide Web its main page, famous for its simple single search bar and Google logo, was remarkably cleaner and had fewer search options, and unlike its competitors no encyclopedia like topical search functions. However its SERP display was not that much different than its competitors. These designs from the 1990’s had evolved out of indexing designs (Singhal, 2001), and displayed results typically with a hyperlinked title, a meta description, if one was available and a site Uniform Resource Locator URL (Glossbrenner, 1999; Hock, 1999). Many also displayed other search options included with each result like ‘more like this’ links or in some cases translating options (Glossbrenner, 1999; Hock 1999). Not all of Google’s competitors used the same order of these SERP result features. Additionally, Google did have a commitment to search results over advertising, limiting adds to only text (Battelle, 2005, p.92). Using Google’s SERP design as a base, GS kept the clean, add free design, and title,
author and publisher information, meta descriptions, number of citations and links to other works (see figure 2.2). GS and has undergone few changes since its release in 2004 (Jacsó, 2008).

Figure 2.2 Google Scholar SERP Design

Prior SERP and Document Triage Methodological Approaches

Experimental methods are common for exploring document triage (Loizides & Buchanan, 2011; Bae et al., 2010; Loizides & Buchanan 2009; Buchanan & Owen, 2008; Buchanan & Loizides, 2007; Badi, 2006; Bae 2005; Shipman et al, 2004; Marshall & Shipman, 1997) but there have been some qualitative studies as well (Bae et al. 2006; Adler et al., 1998; Cool et al., 1993). Experimental designs are also common for SERP studies (Lorigo et al., 2008; Cutrell, & Guan 2007; Guan & Cutrell 2007; Pan et al., 2007; Joachims et al., 2005). Due to this study’s goal to investigate the nature of document triage behavior and its relationship to SERP design and that most studies exploring document triage have been experimental, this study will follow suit with an experimental design.

Repeated Measure Design

Repeated measures design is an often used experimental approach for both SERP and document triage studies. Cutrell and Guan (2007) experimented with increasing the meta descriptions in SERPs with more information about web page documents using a repeated
measures design. Using a repeated measure within subjects design, another study tested how
changes in text size within PDF thumbnail images of articles affected document triage behavior
at the within document triage stage of Loizides (2012) funnel model (Buchanan & Owen, 2008).
Considering how other SERP and document triage studies have used the repeated measures
method, using this same method for this study seemed appropriate and has the benefit of
reducing participant sample size. A smaller sample size does have the problems of the group
potentially being too small to detect real differences and the participant group may not be
representative of larger groups.

Eye Tracking

Eye tracking is a technique that uses special equipment to allow an experimenter to be
able to tell when and where someone is looking and for how long. Before eye tracking
equipment became less expensive Salvucci and Goldberg (2000) developed protocols for
identifying fixations and saccades in eye tracking. Eye tracking is often used in the field of
psychology (Rayner et al, 2009) for many purposes including studying visual encoding and
autonomic activity (Lee et al, 2004). In the field of HCI, Poole and Ball noted in 2006 that eye
tracking had becoming increasingly employed to study usability in HCI contexts. Shortly
thereafter Ehmke and Wilson (2007) wrote a paper about using eye tracking methods for
identifying web usability problems. Every SERP study researched for this proposal (Lorigo et al.,
2008; Cutrell, & Guan 2007; Guan & Cutrell 2007; Pan et al., 2007; Joachims et al., 2005) used
eye tracking methods, as did one of the document triage studies (Loizides & Buchanan, 2011).
Because of how common it is to use eye tracking tools in HCI studies and because of their usage
in SERP studies in particular, it seemed appropriate to use eye tracking tools and methods in this
study as well, allowing the investigation of how the different designs and areas of interest draw
visual attention as well as recording navigation behavior and time on task. Several of the aforementioned studies use eye trackers to investigate areas of interest relating to SERP design (Dumais, Buscher, & Cutrell, 2010; Joachim’s et al., 2005 & Lorgio et al., 2008) telling us areas other than the search results like advertisements, search boxes, and other elements of Google’s SERP had on gaze attention. The findings showed that researchers fall into three groups: those who look at a SERP broadly, others who focus only on the results, and those who regularly look at advertisements. This study will similarly investigate the attention in the SERP title and meta description areas of interest as well as the other elements unique to Google Scholar, such as authors and citation information, and determine how long these areas attract attention.

**Relevancy Tasks**

Many studies have explored how changes in multiple factors, such as interface design (Badi, et al., 2010), number and type of screens (Loizides & Buchanan, 2009) and electronic document vs. physical document (Loizides & Buchanan, 2007) affect perception of relevancy. These studies all use a similar method of asking the participant to play a role of a librarian researching a topic and either ranking a document’s relevancy or sorting them into predefined categories and comparing these results with a subject matter expert (SME). SERP studies assign participants specific questions to answer with the search engine and rate participants if they were able to answer the questions properly or not (Cutrell, & Guan 2007; Pan et al., 2007 & Joachims et al., 2005). For these reasons in this study participants will be asked to make a binary decision as to whether a document is relevant to one predefined category or a second predefined category. This is also one of the four common methods used to assess an information retrieval tool’s accuracy (Saracevic, 2008, p. 769). According to the document triage funnel model, the triage and reading process are components that all lead to an ultimate decision of either accepting or
rejecting a document in this binary fashion, so the studies design should also be binary. This change reduces cognitive load on participants, allowing them to focus more on the task at hand rather than spend time deciding at what level it is relevant. Considering the problem facing students is making relevancy decisions while sifting though documents, comparing their choices with that of a subject matter expert will more exactly inform us if the design change helps alleviate this problem.

One other change in design from the norm for document triage studies is the number of documents that participants will be able to view. Several studies had participants assess 40 documents in their experiments (Bae, et al, 2005; Bae et al. 2006, Bae et al 2010) while one study used 20 (Buchanan and Loizides, 2007). Buchan and Loizides (2007) used only 20 documents in their study. Joachims et al.’s (2005) study found that 96% of students tested only viewed the first page of results. Therefore this study will use only 20 search results in its design.

**Surveys**

Bae et al. (2010) concluded their study with a post-task questionnaire in order to find subjective impressions of VKB’s design, finding that most of the participants indicated that the visualization capabilities were helpful in identifying documents of interest and organizing them. Similarly Dumais, Buscher and Cutrell (2010) use a survey questionnaire to attempt to determine relationships between the identified gaze pattern groups and the subjective impressions of the search engine being tested and their general search behaviors, but most of the differences found were not statistically reliable. This study will also use a post-task survey to report subjective impressions of the SERP design change being tested, and compare participant impressions with other data collected.

**Conclusion**
The fields of philology, psychology, economics and information science have all investigated and found support for the Principle of Least Effort, and important motivational component to document triage. SERPs have been investigated discovering the importance rank ordering has on relevancy perceptions (Lorgio et al. 2008; Pan et al., 2007; Joachems et al., 2005), and the effects elements of SERP other than results have on gaze patterns (Dumais, Buscher & Cutrell, 2010). Cutrell and Guan (2007) found that longer meta descriptions in SERPs improved the accuracy of information tasks. Document triage typically involves information seekers using SERP while seeking relevant documents.

Document triage is the process of quickly finding relevant documents like web pages, periodical articles, and other published materials located using a search engine, automated delivery system or from a human intermediary. Fernando Loizides (2012) built a model for document triage behavior involving IR tools called the ‘funnel model’. This model breaks the document triage process into three parts, the ‘surrogate triage stage’, the ‘within document triage stage’, and the ‘further reading stage’. The ‘surrogate triage stage’ involves reviewing SERPs, and other document surrogates like journal article web pages containing article abstracts and summaries. The ‘within document triage stage’ involves quickly reviewing entire documents rather than their surrogates while making a relevancy decision. The ‘further reading triage stage’ refers to when previously accepted documents are more thoroughly read and evaluated. Most of the relevancy decisions about retrieved documents are made at the ‘surrogate triage stage’ due to the limited depth of triage a user is willing to undertake when looking through a search results list (Spink et al., 2004; Buchanan & Loizides, 2007). Specific document areas such as title, abstract, and section heading text, have been shown to receive the most attention during the

Abstracts are information packets that are written to convey the meaning of a document’s subject content. Because abstracts are an area that receives some of the most visual attention during the document triage process, it is sensible to augment a SERPs meta description with an academic article’s abstract in a search engine that indexes academic documents. Loizides’ (2012) ‘funnel model’ involves SERPs at the ‘surrogate triage stage’ in his model and many of prior SERP studies use tasks that could be described as document triage tasks (Lorgio et al. 2008; Cutrell & Guan, 2007; Joachems et al., 2005 Dumais, Buscher, & Cutrell, 2010 Pan et al., 2007). Relevancy, though a rich term with many debated meanings, can be defined as pertaining to the matter at hand (Saracevic, 1996) and is a key concept involved with the document triage process and purpose of abstracts.

A review of the literature has shown that GS is a search engine that indexes academic scholarly articles, and though it has been studied in other ways, GS has not yet had its SERP investigated. Because of the lack of previous study of the GS SERP and due to it similarity to Google and the likelihood students will not need prior training, GS and its SERPs were selected for this study. In order to investigate GS and its unique SERP elements this study used eye-tracking equipment and methods common to both SERP and document triage studies that have been reviewed.

Based on this review of the literature, this experiment replaced the existing GS SERP meta descriptions with the abstracts of the results corresponding articles and used an experimental repeated measure design to test the effects this design change had on freshmen undergraduate abilities in determining article relevancy. The findings bring new understanding of
how the meta description augmentation affect gaze time in areas of interest as well as how the design change may affect undergraduates document triage behavior.
CHAPTER 3 – METHODOLOGY

The methods section begins with an explanation of this study’s design, followed by a discussion of each of the instruments used in the study. Then the participants recruited and study setting will be briefly described. Next the data collection procedure will be explained in detail, followed by an in-depth description of the proposed data analysis techniques and how they related to this study’s research questions. A discussion of this study’s pilot, its findings, and how those findings affected this proposed study’s design will come next, followed by a discussion of this study’s limitations.

Study Design

This experiment used a repeated measure within factors experimental design to test the effects, if any, that a change in the results design within the Google Scholar’s results list might have on freshmen undergraduate student document triage behavior. A repeated measures design uses the same subjects as both the control and the experimental group. The factor, or variable, tested by each participant was the design change to Google Scholar’s results display’s meta description. The control display was be a list of Google Scholar results generated by a keyword search entered by the researcher that leaves the results meta descriptions unaltered seen in figure 3.1.

![Figure 3.1: Control Google Scholar Result Example](image)

The experimental display was be a list of Google Scholar results generated by a keyword search entered by the researcher that replaces the existing meta descriptions with each individual result’s corresponding abstract as seen in figure 3.2.
Figure 3.2: Experimental Google Scholar Result Example

All of Google Scholar’s search results and documents were accessed and read with the same Internet browser, Internet Explorer, and the screen’s activity was recorded using Tobii Eye tracking software.

In order to prevent the students from looking through the same sets of results twice, the same search string was used to collect 20 total results which have full abstracts used to augment the SERPs meta description. In order to limit possible effects of fatigue, each set of results was limited to 10 entries each. One page of 10 results was designated SERP group A, the other was designated as SERP group B. Both groups were recoded in both the control and experimental design. These control and experimental results pages were coded to ensure that the citation functions and the links to full PDFs included with each result were working properly (see appendix B). These recoded pages were mounted on the University of Missouri’s Bengal server. This ensured that participants were all viewing the same results as they performed the study’s task.

Counterbalancing was used to randomize the order in which the students saw the two different sets of search results, as well as which set of results were displayed in the control or the experimental format. These sets of results will be referred to as ‘SERP group A control’ and ‘SERP group B control’ or ‘SERP group A experimental’ and ‘SERP group B experimental’ depending on the version of the display’s meta description. These counterbalancing measures
should help reduce contamination of the findings due to factors such as fatigue, the quality of abstracts in one set of results compared to the other, and the order of exposure. In order to accomplish this counterbalancing four participant groups were created, randomizing the results display pairing and the order of exposure as shown in table 3.1.

<table>
<thead>
<tr>
<th>Participant Group</th>
<th>Search Results</th>
</tr>
</thead>
</table>
| Group 1           | SERP group A Control  
|                   | SERP group B Experimental                           |
| Group 2           | SERP group A Experimental  
|                   | SERP group B Control                                |
| Group 3           | SERP group B Control  
|                   | SERP group A Experimental                           |
| Group 4           | SERP group B Experimental  
|                   | SERP group A Control                                |

Participants were asked to play the role of a student who is working on a research paper about cyberbullying. Participants were given the task of reviewing a series of ten control search results and their corresponding articles, as well as ten experimental results and their corresponding articles in a randomized fashion as discussed above. As they review the retrieved results they were asked to sort the referenced articles into two categories that they fit best, one category pertaining to how cyberbullying affects people, and the other category relating to cyberbullying prevention and policies. (See Appendix C). As they searched through the articles, participants were asked to mark which results were relevant to which category on a printed list of the titles for the results displayed. The topic was chosen due to potential student interest in the
topic, assuming the likelihood of their awareness of the issue considering most freshmen have just graduated from high school.

The study’s design allowed for a number of dependent variable data sources to be collected and recorded which are listed below:

1. Participant’s relevancy scores
2. Areas of interest (AOI), namely the title, author and publisher information, meta descriptions, and number of citations and links to other works (see figure 3.3).
3. The number of times the participant links to other documents in each set of results
4. Total time required to complete the document triage task within each set of results
5. Participant’s impressions of the two environments, collected along with demographic data, in the post experiment survey

Figure 3.3: Areas of Interest

Instruments
Eye Tracker Instrument

Much of the study’s observational data were recorded using Tobii X2-60 eye tracking equipment and software. The Tobii eye tracker makes use of infrared lights to track and calculate a gaze plot, following a participant’s view of the screen. While the system can record the voice and the image of the users, these data were not collected in this study. Instead, the eye tracker recorded the gaze time in the areas of interest on the SERPs. The whole session was unobtrusively screen recorded along with the eye tracking data, allowing it to be played back for data analysis as necessary. The software recorded how often links on a page were used, how long they were on each page, and the overall time spent on each task. These metrics collected via the Tobii software allowed for the data to be analyzed without having to review files in real time. The Eye Tracker has large head movement tolerance, allowing participants to move freely and naturally and eliminating the need for recalibration during long sessions. It is a highly accurate gaze-position system that captures data at 60 Hz. The dual camera system includes automatic selection of bright or dark pupil eye tracking that follows both. Tobii Analytics SDK 3.0 is the software that was used to store gaze data and analyze user behavior (Tobii Eye tracking research, 2013). The laptop used was a Dell Latitude E6530 with 238 gig of storage, 6 MG of RAM, it was running Windows 7 Enterprise. The laptop had a 15.6" display and an Intel (r) Core (TM) i7-370QM CPU running at 2.70 GHz.

Document Relevancy Instrument

As participants worked through the relevancy assessment task, they were asked to assign each result to one of two categories that they fit best, one pertaining to how cyberbullying affects people and the other relating to cyberbullying prevention and policies. Participants had a printed
list of each of the referenced articles’ titles. Each title was displayed in the same order they are on the SERP. Underneath each title the two categories were printed. Participants were asked to circle which category the articles’ title best fits. The participants’ results were each compared with a subject matter expert Douglas Abrams’ sorting decisions of the same articles and given a percentage accuracy score.

Survey Instrument

The post experiment survey collected demographic information about the freshmen undergraduates including age, gender, the approximate number of classes they have taken, the size of their high school graduating class and number of college credit hours they have completed. The survey asked participants about their familiarity with Google Scholar and the topic of cyberbullying for reporting purposes. Considering how user perceptions and preferences can affect how useful participant’s find a design (Nielson & Levy, 1994), the remaining questions on the survey asked participants if they noticed any differences in the time it took them to complete the tasks, which design did they prefer, which one they would rather use again, how confident they felt using the different designs, and if they remembered opening more web pages in one design environment or the other (see appendix D).

Participants

Using G*Power 3.1, a statistics estimator available from the Universitat Dusseldorf (web page http://www.gpower.hhu.de/en.html), it was calculated that for an effect size F of .35 at an alpha level of .05 the study would require 32 freshmen undergraduate participants, 8 per counter balancing group. Participants were recruited verbally from pedestrian student traffic in front of London hall and at Speaker’s Circle. Participants were offered a $10 USD incentive for taking part in the study. They were screened for Tobii eye tracking requirements, as the eye tracker has
mechanical difficulty with bifocals, trifocals, and progressives, as well as corneal, cataract, or intraocular implants surgeries, and conditions like, lazy eye, strabismus, and nystagmus.

**Study Setting**

The study took place at the Information Experience Laboratory located in 111 London Hall University of Missouri campus in Columbia MO. The room was furnished with padded chairs and has two desks, one of which was set up with a laptop computer fitted with eye-tracking hardware.

**Data Collection Procedure**

Participants were scheduled to meet with the researcher at the Information Experience Laboratory. After signing the consent form, each was given a brief description of the study, how the eye tracker works, and an explanation of what to expect. Participants were taken through the eye tracker calibration process. The participants were asked to play the role of a student doing research for a class paper, with the information needs outlined in their task explanation and reference sheet. Their goal was to review two sets of ten search results and mark which results were relevant to certain categories on printed sheets containing the search result titles and authors that were be given to them by the researcher.

Once the participant explained that they understood their task and what to do, they were left alone in the room by the researcher, having been instructed to follow a screen prompt to start the experiment. Once started, the Tobii eye track software then displayed the first set of results appropriate to the randomly assigned counterbalancing group for each participant. The participant was then free to explore the results and corresponding pages as they would normally. After finishing reviewing the first set of results and recording which results they found relevant to their task, they were prompted to hit a designated F10 command key to end the first task and
start the second. The command then opened the next set of ten search results appropriate for their group. Once finished with the second task they were prompted to press the same command press key to end the study and then prompted to complete a brief post experiment survey. After finishing both tasks and the survey the participant was instructed to inform the researcher, who was in the adjacent room and available for questions in case the participant ran into any problems. A participant session typically lasted 30-40 minutes.

**Data Analysis**

An analysis of all the collected data, the mean gaze time in areas of interest, mean triage stage transitions, mean total time on task, and mean relevancy scores combined answered this study’s first research question: how does augmenting the meta description with an abstract in a SERP effect document triage behavior in undergraduates? (See table 3.2)

<table>
<thead>
<tr>
<th>Research Questions</th>
<th>Data Sources</th>
<th>Analysis methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>How does replacing the meta description with an abstract in a SERP effect document triage behavior in undergraduates?</td>
<td>gaze time in areas of interest, mean triage stage transitions, mean total time on task, mean relevancy scores, survey data</td>
<td>Paired t-test, analysis of variance</td>
</tr>
<tr>
<td>How does replacing the meta description with an abstract in a SERP effect document triage behavior in undergraduates at the surrogate triage stage?</td>
<td>gaze time in areas of interest, and mean triage stage transitions, survey data</td>
<td>Paired t-test, analysis of variance</td>
</tr>
<tr>
<td>How does replacing the meta description with a SERP effect perception of document relevancy?</td>
<td>relevancy scores, survey data</td>
<td>Paired t-test, analysis of variance</td>
</tr>
<tr>
<td>How does augmenting the meta description with abstracts in a</td>
<td>mean total time on task, mean relevancy scores, mean triage</td>
<td>Pearson correlation coefficient, paired t-test</td>
</tr>
</tbody>
</table>
Gaze time in areas of interest, and mean triage stage transitions data were used to answer the second research question: how does augmenting the meta description with an abstract in a SERP affect document triage behavior in undergraduates at the surrogate triage stage? (See table 3.2)

Mean relevancy scores compared to that of a subject matter expert in each results set were used to answer the third research question: how does augmenting the meta description with a SERP affect perception of document relevancy? (See table 3.2)

The mean total time on task in each set of documents along with combined analysis of the accuracy of the participant’s relevancy scores, the mean triage stage transitions, and gaze time in areas of interest were used to answer the fourth and final research question: how does augmenting meta description with abstracts in a SERP affect document triage efficiency? For the purposes of this study, efficiency will be defined as the ratio of time to accuracy of the relevancy-sorting task. A post-task survey collected demographic data and subject impression of the SERP design change and how it affected relevancy decision and document triage behavior.

Following are descriptions of how each data variable was analyzed.

**Areas of Interest Gaze Time**

The eye tracker recorded the amount of participant gaze time within four groups of areas of interest: the title, the meta description, the author/publisher information and the cited element. The mean amount of time spent in each set of areas of interest groups were compared between both the control and experimental set of search results. These data were tested with the a paired t-test, and an analysis of variance.
The specific hypotheses to be tested are shown here in the null form:

1. \( H_0 \): There is no significant relationship between the inclusion of abstracts within Google Scholar SERP and the amount of time titles are viewed.

2. \( H_0 \): There is no significant relationship between the inclusion of abstracts within Google Scholar SERP and the amount of time meta descriptions are viewed.

**Triage Stage Transitions**

The eye tracker creates a record of all the page navigations which occur while tracking a web page. Because of this, all the number of transitions from the results lists to other web pages and transitions from surrogate triage stage to within document triage stage were recorded. Both the mean number of transitions between the set of search results and other web pages and the mean number of transitions between the surrogate triage stage and the within document stage were compared between both the control and experimental set of search results. These data were tested with a paired t-test, and an analysis of variance. The specific hypotheses to be tested are shown here in the null form:

3. \( H_0 \): There is no significant relationship between the inclusion of abstracts within Google Scholar and the median number of transitions between surrogate triage stage documents.

4. \( H_0 \): There is no significant relationship between the inclusion of abstracts within Google Scholar and the median number of transitions between within document triage stage documents.

**Total Time on Task**

The eye tracker also creates a record of the total time on task. The mean amount of time on task was compared between both the control and experimental sets of search results. These
data were tested with a t-test, and an analysis of variance. The specific hypotheses to be tested are shown here in the null form:

5. \( H_0 \): There is no significant relationship between the inclusion of abstracts within Google Scholar and the total amount of time required to complete the research task.

Relevancy Scores

The recorded relevancy scores were compared against those of a subject matter expert generating a percentage accuracy score. The mean amount of time on task was compared between both the control and experimental sets of search results. These data were tested with a paired t-test, and an analysis of variance. The specific hypothesis to be tested are shown here in the null form:

6. \( H_0 \): There is no significant relationship between the inclusion of abstracts within Google Scholar and the scores for relevancy accuracy.

All hypotheses were tested at a minimum of the .05 level of significance and calculated using IBM’s SPSS software.

Survey

The post-experimental survey consisted of seven demographic questions. The first and second questions asked participant age and the second gender for reporting purposes. The third, fourth, and fifth questions helped ascertain relevant freshman undergraduate research experience, asking how many papers participants had written for class, what this size of their graduating class was and how many credit hours they had completed respectively. The sixth and seventh survey questions asked participants to self-assess their familiarity with Google Scholar and cyberbullying respectively. All of the prior research experience, GS, and cyberbullying familiarity questions were checked against the total time on task and relevancy scores in both the
control and experimental designs categories to see if a linear regression could be done to discern their effects on these elements of the experiment.

The final five research questions all relate to the research questions. Survey question number eight; did you feel like you were faster in one version of Google Scholar or another, was used to answer the first and fourth research questions.

RQ 1: How does augmenting the meta description with an abstract in a SERP affect document triage behavior in undergraduates?

RQ 4: How does replacing the meta description with abstracts in a SERP affect document triage efficiency?

Survey question number nine; of the two versions of Google Scholar did you like one better than the other, was used to answer the third and fourth research questions.

RQ 3: How does replacing the meta description with a SERP affect perception of document relevancy?

RQ 4: How does replacing the meta description with abstracts in a SERP affect document triage efficiency?

The tenth survey question; if you had to use Google Scholar for a class research paper would you prefer one version over the other, was be used to answer the third and fourth research questions.

RQ 3: How does replacing the meta description with a SERP affect perception of document relevancy?

RQ 4: How does replacing the meta description with abstracts in a SERP affect document triage efficiency?

The eleventh survey question; did you remember feeling more confident with one version of Google Scholar than the other, was used to answer research question number three.
RQ 3: How does replacing the meta description with a SERP affect perception of document relevancy?

The twelfth and final survey question, do you remember opening other webpages or PDF documents more often in one environment than the other, was used to answer research questions 1 and 2.

RQ 1: How does augmenting the meta description with an abstract in a SERP affect document triage behavior in undergraduates?

RQ 2: How does replacing the meta description with an abstract in a SERP affect document triage behavior in undergraduates at the surrogate triage stage?

**Findings From the Pilot**

**Methodology**

The pilot for this study also tested the effects of augmenting Google Scholar SERP meta descriptions with abstracts using a repeated measures within samples experiment. The pilot consisted of 8 participants with master’s level education or higher and used counter-balancing procedures to break participants into 4 groups of 2. The participants’ task was to assess 40 results total, 20 relating to tea and caffeine effects and 20 relating to tea and caffeine effects and rank them for relevancy. Depending upon the counterbalancing group, participants encountered either the coffee topic or the tea topic first in either the experimental design, containing the abstract augmented meta description SERP, or the control, containing an unaltered meta description SERP (see table 3.3).

<table>
<thead>
<tr>
<th>Participant Group</th>
<th>Search Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>Coffee Control</td>
</tr>
</tbody>
</table>
Participants were asked to play the role of a student who is working on a research paper about the effects caffeine from tea and coffee have on the human body. They were given guidelines to seek articles relating to how coffee and caffeine and tea and caffeine affect the human body. When selecting which articles were the most relevant to the topic of the effects coffee and caffeine have on the human body, the first priority was to find harmful effects, followed by the benefits caffeine has on the human body.

The subject was chosen because both beverages are popular and there may be student interest in the topic, due to how broadly available and consumed both coffee and tea are in the US. As they searched through the articles, participants were given a printed list of the titles for each of the results displayed and were asked to mark which 5 results were the most relevant to needs provided in the guidelines with a + and the 5 that were the least relevant with a -. The results marked with a + were scored with 2 points, those blank 1 point and those with a – 0 points for the sake of ranking the articles’ perceived relevancy. Areas of Interest were not studied in the pilot. The pilot did record the total time on task in each SERP design and the number of transitions to other surrogate document pages. The pilot also contained a brief post-experimental interview, asking participants about their experiences determining article relevancy in the two different SERP designs.
Results

The pilot study found that the differences in total time on task between both the experimental and control groups were not significant: Paired samples T-test $T(7)=.034$, $p > .05$, $d = 9.33$ ANOVA $(F(1,14)= .017$, $p.900$. The pilot also showed that transitions to other Web pages did change significantly between the control and the experimental groups: Paired samples T-test $T(7)=3.306$, $p < .05$, $d = 6.52$. ANOVA $(F(1,14)= .017$, $p.900$.

Post-experiment interview results found that participants noticed that they made fewer clicks in the experimental environment compared to the control. All 8 of the participants also preferred the experimental environment to the control environment. Participants did report that completing the task took less time in the experimental SERP design, but this was a false perception as the recorded time on task showed no meaningful change in total time on task between the two groups. Participants also reported titles alone were more confusing and less informative and that having the abstracts made it easier to compare articles. One participant did report that they thought that by having the abstract it could make overall investigating less thorough.

Perceptions of relevancy shifted depending upon which SERP design was being used. Within the coffee SERPs results 4, 7, and 10’s rank shifted by more three average points or more (see figure 3.4).
The tea SERPs had results 1, 2, and 9’s ranking shift by 3 or more average points depending on which SERP environment they were viewed in (see figure 3.5).
Discussion

These findings show that augmenting Google Scholar’s meta description with abstracts does effect how results relevancy is perceived. It also showed that even though mean time on task did not change from the control group to the experimental group, participants thought they were triaging faster when using the experimental design. All of the participants preferred the experimental design. Also the inclusion of abstracts changed participant document triage behavior by reducing how often the participants navigated away from the SERP to seek more information from other document surrogates or documents themselves. The design change did create an environment that required less effort from the participants to determine relevancy, but without a subject matter expert to compare ranking results it is hard to say if the change in perceived relevancy was a beneficial one or not.

Lessons Learned

The pilot also revealed problems with the data collection design. The study’s retrieved search results were predominately from medical journals and both the abstracts and unaltered meta descriptions were highly technical. Determining which articles were relevant to the task required a high degree of subject matter expertise and likely affected the study’s results. The scaling of article relevancy made the task more cognitively difficult for participants and was not as useful as comparing participant results with the results of a subject matter expert could have been. For these reasons a different subject for the task was used, as well as a different method for determining result relevancy.

Another improvement the pilot suggested for the proposed study was to reduce the number of results from 40 total to 20. The total time on task results showed that 5 out of the 8 participants spent less time in the second SERP design environment regardless of if it was the
control design or the experimental design environment. Another rationale for this design change is that not one of the average relevancy scores for the last ten results for either coffee or tea subjects in either the control or experimental design environment changed more than 2 points. These results suggest that the last 10 results are receiving less attention and that participants are becoming fatigued as they look through and assess so many articles. Rather than using experienced researchers, one final change from the pilot was to use freshmen undergraduates, for the reasons already discussed in the introduction.

**Limitations**

This experiment had several limitations that need to be discussed. One is that the search task was not a natural one, but one contrived for this study. Participants were asked to evaluate all 10 results on each page they viewed. This is also not usual information seeking behavior. Typically, only the top results are reviewed before a new search query is created. Participants may or may not have had much interest in the search topic. The researchers did not generate the search query for the relevancy task. A more naturalistic study would allow participants to create their own search strings and investigate their own topics.

Eye tracking research has several base limitations. Eye tracking does not provide certainty that users saw something conscientiously, because users can aim their eyes at an area for a short period of time without any awareness. Eye tracking does not inform us that users did not see something, since eye tracking does not capture peripheral vision. Eye tracking cannot tell you why users are looking at something. Eye tracking cannot test everybody effectively, because problems can occur when some users wear eyeglasses or hard contacts, have small pupils, a wandering eye or an overly expressive face (Nielson & Pernice, 2010). A limitation specific to this study was that in order for Tobii to function the screen resolution had to be high and the font
sizes 11pt or smaller, otherwise fixation points would be too large to point with any precision. While not a problem for this user group, the small font size limitation could have been a hindrance for anyone requiring a larger font.

The study’s design also has no real way to assess different skill levels and assumed that the freshmen were all equally skilled in searching and had a similar knowledge level. Considering that the pilot found that the design change did cause a shift change in how some documents were ranked in relevancy, skill is not a requirement of this experiment; the design change should affect both the skilled and the unskilled. This assumption will limit how much effect the design is truly having on the relevancy task scores and total time on task. Finally, there is a problem in that SERP designs are too different from one another for a GS display study’s findings to be generalizable to all other SERP designs. The review of the literature revealed that abstract quality is a key component for abstracts to be of use; this study has no way to assess abstract quality and only uses counterbalancing to help reduce the effect the differing quality of abstracts may have on the relevancy scoring task. There are also mathematical limitations in that the quantitative measure assumes that there will be a standard deviation.

**Conclusion**

This study employed a repeated measure within factors experimental design. The independent variable was a change in the meta-description within Google Scholar, an information retrieval tool. The independent variable was a change in an information retrieval tool, Google Scholar’s, meta-description. The study needed a sample of 32 undergraduate freshman for an effect size of .35 at an alpha level of .05 according to the G*Power 3.1 statistics estimator. All of the data was collected in the University of Missouri’s Information Experience laboratory located in London Hall.
Undergraduate participants were asked to view two different Search Engine Results Pages (SERP). They were then free to navigate these as they would normally and determined which resources were relevant to their assigned research task. One SERP was in control format and the other in the experimental format. The eye tracking, recorded screen data, and relevancy scores allowed a number of dependent variable data sources to be collected. Together these data helped answer the study’s first and overarching research question: how does augmenting the meta description with an abstract in a SERP affect document triage behavior in undergraduates?

The experiment recorded differences in participant time looking at areas of interest, mean time on task, the survey instrument, and the number of links to other documents. These data helped in answering the second research question: how does augmenting the meta description with an abstract in a SERP affect document triage behavior in undergraduates at the surrogate triage stage? The accuracy of the participant’s relevancy scores compared to that of a subject matter expert in each results set and the survey instrument answered the third research question: how does augmenting the meta description with abstracts in a SERP affect perception of document relevancy? The fourth and final research question, how does augmenting the meta description with abstracts in a SERP affect document triage efficiency, was answered by recording total time on task in each set of documents, the survey instrument and with combined analysis of the accuracy of the participant’s relevancy and the number of links to other documents. These findings lead to more insight as to how SERP design can improve document triage in the hope helping to help alleviate some of the problems freshmen undergraduates face in making sense of the ever-growing amounts of information available to students today.
CHAPTER 4 - RESULTS

The purpose of this analysis is to determine the effects of adding articles’ abstracts to GS’s SERP design have on undergraduate freshmen’s document triage behavior through three primary data sources collected during the experiment: the data collected by the eye tracker, the participant relevancy responses, and the post experiment survey. A total of 32 participants completed the relevancy sorting tasks and provided eye tracking data. The eye tracking data bring new understanding of how the design change affects gaze time in areas of interest; namely the title, author and publisher information, the meta description, and the number of citations and links to other works. The number of document transitions, time on task, and relevancy responses, help shed some light on how the design change affected undergraduates document triage behavior and their perception of relevancy. The survey data report the demographics of the participant population sampled and their subjective experiences of the two different SERP design environments, of the 32 participants only 30 completed the survey.

Demographic Information

The first and second survey questions asked for the participant’s age and gender for reporting purposes shown in table 4.1.

Table 4.1

<table>
<thead>
<tr>
<th>Question 1: How old are you?</th>
<th>Answer</th>
<th>Response</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>17-18</td>
<td>2</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>18-19</td>
<td>19</td>
<td>63%</td>
<td></td>
</tr>
<tr>
<td>20-21</td>
<td>4</td>
<td>13%</td>
<td></td>
</tr>
<tr>
<td>21-older</td>
<td>5</td>
<td>17%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Question 2: What is your gender?</th>
<th>Answer</th>
<th>Response</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>17</td>
<td>57%</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>13</td>
<td>43%</td>
<td></td>
</tr>
</tbody>
</table>
The third, fourth, and fifth survey questions helped ascertain relevant freshman undergraduate research experience. The responses to question three and four are seen in as seen table 4.2.

Table 4.2

*Question 3: Approximately (best estimate) how papers have you written for your college classes?*

<table>
<thead>
<tr>
<th>Answer</th>
<th>Response</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-6</td>
<td>7</td>
<td>23%</td>
</tr>
<tr>
<td>7-9</td>
<td>11</td>
<td>37%</td>
</tr>
<tr>
<td>10-12</td>
<td>4</td>
<td>13%</td>
</tr>
<tr>
<td>13-15</td>
<td>3</td>
<td>10%</td>
</tr>
<tr>
<td>16 or more</td>
<td>5</td>
<td>17%</td>
</tr>
</tbody>
</table>

*Question 4: How many people where in your graduating class?*

<table>
<thead>
<tr>
<th>Answer</th>
<th>Response</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than 100</td>
<td>4</td>
<td>12%</td>
</tr>
<tr>
<td>101-300</td>
<td>6</td>
<td>21%</td>
</tr>
<tr>
<td>301-500</td>
<td>10</td>
<td>34%</td>
</tr>
<tr>
<td>501-1000</td>
<td>8</td>
<td>28%</td>
</tr>
<tr>
<td>1000 or more</td>
<td>1</td>
<td>3%</td>
</tr>
</tbody>
</table>

Participants were told verbally at the beginning of each session they could report applied credit hours from high school as well as those earned at the university. These results can be seen in table 4.3.

Table 4.3

*Question 5: Approximately (best estimate) how many credit hours have you completed, (not including this semester)?*

<table>
<thead>
<tr>
<th>Answer</th>
<th>Response</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than 12</td>
<td>1</td>
<td>3%</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>3%</td>
</tr>
<tr>
<td>15</td>
<td>10</td>
<td>33%</td>
</tr>
<tr>
<td>18</td>
<td>5</td>
<td>17%</td>
</tr>
</tbody>
</table>
The sixth and seventh survey questions asked participants to self-assess their familiarity with GS and the topic of cyberbullying respectively. These data are shown in table 4.4.

Table 4.4

**Question 6: How familiar are you with Google Scholar?**

<table>
<thead>
<tr>
<th>Answer</th>
<th>Response</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>I have never used Google Scholar before.</td>
<td>11</td>
<td>37%</td>
</tr>
<tr>
<td>I have used Google Scholar once or twice.</td>
<td>11</td>
<td>37%</td>
</tr>
<tr>
<td>I have used Google Scholar several times before.</td>
<td>7</td>
<td>23%</td>
</tr>
<tr>
<td>I use Google Scholar often.</td>
<td>1</td>
<td>3%</td>
</tr>
<tr>
<td>I almost always use Google Scholar.</td>
<td>0</td>
<td>0%</td>
</tr>
</tbody>
</table>

**Question 7: Did you already know a lot about the topic of cyberbullying?**

<table>
<thead>
<tr>
<th>Answer</th>
<th>Response</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>I have heard of it, but I have never paid much attention to it.</td>
<td>2</td>
<td>7%</td>
</tr>
<tr>
<td>I know a little about it, but I don’t know more than the basics.</td>
<td>4</td>
<td>13%</td>
</tr>
<tr>
<td>I know some about it from the news and classes, enough that I could explain it to some one else.</td>
<td>18</td>
<td>60%</td>
</tr>
<tr>
<td>I know more than most people from the news, classes, and my own interest in the topic.</td>
<td>5</td>
<td>17%</td>
</tr>
<tr>
<td>I know a lot about the topic, I have read articles, watched videos, or been to lectures about it.</td>
<td>1</td>
<td>3%</td>
</tr>
</tbody>
</table>

All of the prior research experience, GS, and cyberbullying familiarity questions were checked against the total time on task and relevancy scores in both the control and experimental deigns.
categories to see if a linear regression could be done to discern their effects on these elements of the experiment. None of the results passed the assumptions of linearity or showed a spearman’s rank-order correlation except for participant’s response to the GS familiarity question, which had a weak relationship between total time on task in the control condition. But the heteroscedasticity of these GS familiarity data caused them to fail the assumptions necessary for a linear regression. These scatterplots Spearman’s r results are shown in appendix E.

**Heat Maps**

Eye tracking is a technology that tracks the exact point at which a user’s gaze is fixated on a screen. This technology allows how much users looked at different parts of a Web page to be tracked and shown. Following are the heat maps for each of the four counterbalancing participant groups. Figure 4.1 displays aggregate heat maps for the control groups and figure 4.2 displays aggregate heat maps for the experimental groups. Areas where users looked the most are colored red; the yellow areas indicate fewer fixations, followed by the least-viewed green areas. Colorless areas didn't attract any fixations.
While these heat map measures provide some indication of the quality of the results, they do little to improve our understanding of how searchers interact with results. The research reported throughout these results use gaze tracking to enable us to understand detailed patterns of user attention to particular areas of interest.
Figure 4.2 Heat Maps for the Experimental Display for Participant Groups 1-4

Areas of Interest Gaze Time

The Tobii eye tracking software allowed for four group areas of interest to be designated on the GS SERP and the mean participant gaze time to be calculated in each of these group areas of interest per visit to that AOI. This metric measures the duration of each individual visit within an AOI group. A visit is defined as the interval of time between the first fixation on the AOI and the next fixation outside the AOI. Appendix F shows the mean time spent looking at titles per visit in both the control and the experimental environments for each participant. Appendix G shows the mean time per visit spent looking at the author/publisher information in both the control and the experimental environments for each participant. Appendix H shows the mean time spent looking at the meta description per visit information in both the control and the experimental environments for each participant. Finally, appendix I shows the mean time per
visit spent looking at the cited information in both the control and the experimental environments for each participant. These AOI gaze time data help to answer the following research questions:

RQ1: How does augmenting the meta description with an abstract in a SERP affect document triage behavior in undergraduates?

RQ2: How does augmenting the meta description with an abstract in a SERP affect document triage behavior in undergraduates at the surrogate triage stage?

RQ4: How does augmenting the meta description with abstracts in a SERP affect document triage efficiency?

The specific hypotheses tested with these data are shown here in the null form:

1. H₀: There is no significant relationship between the inclusion of abstracts within Google Scholar SERP and the amount of time titles are viewed.

2. H₀: There is no significant relationship between the inclusion of abstracts within Google Scholar SERP and the amount of time meta descriptions are viewed.

Title AOI

A paired-samples t-test was conducted to compare the mean gaze time per visit in the group AOI title in the SERP control and SERP experimental conditions. There was no significant difference in the mean gaze time in the group AOI title gaze time per visit for SERP control (M=00:20.40 SD=00:20.068) and SERP experimental conditions (M=00:15.36, SD=00:11.841) conditions; t(31)=1.839 p =.076. A one-way within subjects (or repeated measures) ANOVA was also conducted to compare the SERP design on group AOI title gaze time. This also showed there was not a significant effect of the SERP design, Wilks’ Lambda =.902, F (1,31) = 3.381, p =.076 at p<.05 level for the two conditions.

Author/Publisher AOI
A paired-samples t-test was conducted to compare the mean gaze time per visit in the group AOI author/publisher in the SERP control and SERP experimental conditions. There was no significant difference in the mean gaze time in the group AOI author/publisher gaze time per visit for SERP control (M=00:08.63 SD=00:08.224) and SERP experimental conditions (M=00:07.80, SD=00:08.577) conditions; t(31)=.667, p =.510. A one-way within subjects (or repeated measures) ANOVA was also conducted to compare the SERP design on group AOI author/publisher gaze time per visit. This also showed there was a not a significant effect of the SERP design, Wilks’ Lambda =.986, F (1,31) = .445, p =.510 at p<.05 level for the two conditions.

**Meta Description AOI**

A paired-samples t-test was conducted to compare the document transitions between the mean gaze time per visit in the group AOI meta description in the SERP control and SERP experimental conditions. There was a significant difference in the mean gaze time per visit in the group AOI meta description scores for SERP control (M=00:27.98 SD=00:25.570) and SERP experimental conditions (M=01:44.62, SD=01:40.728) conditions; t(31)=-5.048 , p =.000. A one-way within subjects (or repeated measures) ANOVA was also conducted to compare the SERP design on group AOI meta description gaze time per visit. This also showed there was a significant effect of the SERP design, Wilks’ Lambda =.549, F (1,31) = 25.485, p =.000 at p<.05 level for the two conditions.

**Cited AOI**

A paired-samples t-test was conducted to compare the mean gaze time per visit in the group AOI cited in the SERP control and SERP experimental conditions. There was not a significant difference in the mean gaze time per visit in the group AOI cited scores for SERP
control (M=00:05.19 SD=00:04.959) and SERP experimental conditions (M=00:03.87, SD=00:03.281) conditions; t(31)=1.353, p =.186. A one-way within subjects (or repeated measures) ANOVA was also conducted to compare the SERP design on group AOI cited gaze time per visit. This also showed there was a not a significant effect of the SERP design, Wilks’ Lambda =.944, F (1,31) = 1.830, p =.186 at p<.05 level for the two conditions.

**Document Transitions**

The Tobii eye tracking software allows for screen recording of each session and logs any Web searching done during the session. Through a review of the screen recordings and Web logs the number of transitions away from the SERP was recorded and tabulated. Appendix J shows the number of surrogate document transitions in both the control and the experimental environments for each participant. These document transition data help to answer the following research questions:

RQ1: How does augmenting the meta description with an abstract in a SERP affect document triage behavior in undergraduates?

RQ2: How does augmenting the meta description with an abstract in a SERP affect document triage behavior in undergraduates at the surrogate triage stage?

RQ4: How does augmenting the meta description with abstracts in a SERP affect document triage efficiency?

The specific hypotheses tested with these data are shown here in the null form:

3. H₀: There is no significant relationship between the inclusion of abstracts within Google Scholar and the median number of transitions between surrogate triage stage documents.
4. H0: There is no significant relationship between the inclusion of abstracts within Google Scholar and the median number of transitions between within document triage stage documents.

A paired-samples t-test was conducted to compare the number of document transitions between the SERP and the database provided webpages in the SERP control and SERP experimental conditions. There was a significant difference in the scores for SERP control (M=6.84, SD=3.878) and SERP experimental conditions (M=4.81, SD=4.306) conditions; t(31)=3.219, p =.003. A one-way within subjects (or repeated measures) ANOVA was also conducted to compare the SERP design on the number of document transitions. This also showed there was a significant effect of the SERP design, Wilks’ Lambda =.749, F (1,31) = 10.363, p =.003 at p<.05 level for the two conditions.

A paired-samples t-test was conducted to compare the number of document transitions between the SERP and the full text documents in the SERP control and SERP experimental conditions. There was not a significant difference in the scores for SERP control (M=6.8438, SD=1.394) and SERP experimental conditions (M=6.56, SD=9.71) conditions; t(31)=1.193, p =.263. A one-way within subjects (or repeated measures) ANOVA was also conducted to compare the SERP design on the number of document transitions. This also showed there was not a significant effect of the SERP design, Wilks’ Lambda =.960, F (1,31) = 1.298, p =.263 at p<.05 level for the two conditions.

**Total Time on Task**

The Tobii eye tracking software allows for screen recording and the timing of a session. Through a review of the screen recordings and time logs the number of transitions away from the SERP was recorded and tabulated. Appendix K shows the total time on task in both the control
and the experimental environments for each participant. These document transition data help to answer the following research questions:

RQ1: How does augmenting the meta description with an abstract in a SERP affect document triage behavior in undergraduates?

RQ4: How does augmenting the meta description with abstracts in a SERP affect document triage efficiency?

The specific hypotheses tested with these data is shown here in the null form:
5. $H_0$: There is no significant relationship between the inclusion of abstracts within Google Scholar and the total amount of time required to complete the research task.

A paired-samples t-test was conducted to compare the total time on task in the SERP control and SERP experimental conditions. There was not a significant difference in the total time on task scores for SERP control ($M=7:37.88$, $SD=4:30.122$) and SERP experimental conditions ($M=7:12.54$, $SD=2:50.458$) conditions; $t(31)=.639$, $p =.527$. A one-way within subjects (or repeated measures) ANOVA was also conducted to compare the SERP design on the total time on task. This also showed there was not a significant effect of the SERP design, Wilks’ Lambda $=.987$, $F (1,31) = 10.363$, $p =.527$ at $p<.05$ level for the two conditions.

**Relevancy Responses**

Each participant filled out a relevancy-sorting instrument while completing his or her assigned document triage task. The recorded relevancy scores were compared against those of a subject matter expert generating a percentage accuracy score. Appendix L shows the relevancy accuracy scores in both the control and the experimental environments for each participant. These document transition data help to answer the following research questions:
RQ1: How does augmenting the meta description with an abstract in a SERP affect document triage behavior in undergraduates?

RQ3: How does augmenting the meta description with abstracts in a SERP affect perception of document relevancy?

RQ4: How does augmenting the meta description with abstracts in a SERP affect document triage efficiency?

The specific hypotheses tested with these data are shown here in the null form:

6. $H_0$: There is no significant relationship between the inclusion of abstracts within Google Scholar and the scores for relevancy accuracy.

A paired-samples t-test was conducted to compare the relevancy responses’ accuracy compared to a subject matter expert in the SERP control and SERP experimental conditions. There was not a significant difference in the relevancy responses accuracy scores for SERP control ($M=83.44$, $SD=15.783$) and SERP experimental conditions ($M=82.81$, $SD=11.705$) conditions; $t(31)=.329$, $p =.745$. A one-way within subjects (or repeated measures) ANOVA was also conducted to compare the SERP design on the number of relevancy responses. This also showed there was not a significant effect of the SERP design, Wilks’ Lambda =.977, $F (1,31) = .108$, $p =.745$ at $p<.05$ level for the two conditions.

**Survey Responses**

Of the 32 participants, two did not fill out the survey despite being asked by the researcher at the end of the session if they had done so. The post experimental survey consisted of seven questions relating to demographics or prior research experience. The final five survey questions all relate to the research questions. All of the survey responses are shown in table
format in appendix M. The first of the final five survey questions, question number eight, helps answer the first and fourth research questions.

RQ 1: How does augmenting the meta description with an abstract in a SERP affect document triage behavior in undergraduates?

RQ 4: How does replacing the meta description with abstracts in a SERP affect document triage efficiency?

This survey question asked: did you feel like you were faster in one version of Google Scholar or another? 6, 20%, responded accurately that neither was faster. The remaining 70% of participants thought that one version or the other made them faster with 9, 30%, responding that they felt faster in the version with shorter descriptions and 15, 50%, reporting that they felt faster in the version with longer descriptions. These three groups, neither was faster, the shorter was faster, and the longer was faster were compared to their actual performances as shown in table 4.5

<table>
<thead>
<tr>
<th>Answer</th>
<th># Perceived Being Faster</th>
<th># of Correct Perceptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neither was faster.</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>I felt faster in the version with shorter descriptions.</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>I felt faster in the version with longer descriptions.</td>
<td>15</td>
<td>6</td>
</tr>
</tbody>
</table>

Survey question number nine helps answer the third and fourth research questions.

RQ 3: How does replacing the meta description with a SERP affect perception of document relevancy?

RQ 4: How does replacing the meta description with abstracts in a SERP affect document triage efficiency?
This question asked, of the two versions of Google Scholar did you like one better than the other? Most, 63%, preferred the SERP design with abstracts, 17% preferred the shorter and the final 20% had no preference either way.

The tenth survey question helps answer the third and fourth research questions. These responses are shown in table 4.3.

RQ 3: How does replacing the meta description with a SERP affect perception of document relevancy?

RQ 4: How does replacing the meta description with abstracts in a SERP affect document triage efficiency?

When asked if they had to use GS for a research paper, the numbers changed slightly at the suggestion of more work. For this question 7% remained undecided, 23% preferred the shorter description, and now 70% preferred the longer meta description.

The eleventh survey question; was used to answer research question number three.

RQ 3: How does replacing the meta description with a SERP affect perception of document relevancy?

This survey question asked if the participants remembered feeling more confident with one version of Google Scholar than the other. 37% didn’t remember feeling more confident with either version while 13% felt more confident with the shorter descriptions. The final 50% felt more confident with the longer descriptions. These two groups, ‘I felt more confident with the shorter descriptions’ and ‘I felt more confident with the longer descriptions’ were compared to their actual performances as shown in table 4.6

Table 4.6
Confidence: Correct perceptions vs. total perceptions

<table>
<thead>
<tr>
<th>Answer</th>
<th># Felt more confident</th>
<th># of Correct Perceptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>I felt more confident with the</td>
<td>4</td>
<td>1</td>
</tr>
</tbody>
</table>
shorter descriptions.

I felt more confident with the longer descriptions.

The twelfth and final survey question helps to answer research questions 1 and 2.

RQ 1: How does augmenting the meta description with an abstract in a SERP affect document triage behavior in undergraduates?

RQ 2: How does replacing the meta description with an abstract in a SERP affect document triage behavior in undergraduates at the surrogate triage stage?

This question asked: do you remember opening other webpages or PDF documents more often in one environment than the other? More participants were able to more accurately access the mean value of document transitions with 17, 57%, reporting I opened more with the shorter descriptions. 8, 27%, reported that they didn't remember a difference with either version. 5,17%, reported that they opened more with the longer descriptions when in actuality this was only the case twice. These two groups, ‘I opened more with the shorter descriptions’ and ‘I opened more with the longer descriptions.’ were compared to their actual performances as shown in table 4.7

<table>
<thead>
<tr>
<th>Table 4.7</th>
<th>Document Transitions: Correct perceptions vs. total perceptions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Answer</td>
</tr>
<tr>
<td></td>
<td>I opened more with the shorter descriptions.</td>
</tr>
<tr>
<td></td>
<td>I opened more with the longer descriptions.</td>
</tr>
</tbody>
</table>

Conclusion

This chapter reported the results of this study’s experiment for the purpose of determining the effects of adding articles’ abstracts to GS’s SERP design have on undergraduate
freshmen’s document triage behavior through three primary data sources collected during the experiment: the data collected by the eye tracker, the participant relevancy responses, and the post experiment survey. The eye tracking data bring new understanding of how the design change affects gaze time in specific areas of interest; namely the title, author and publisher information, the meta description, and the number of citations and links to other works. Of these AOI data the only group AOI that showed a significant difference in gaze time due to the design change was the meta descriptions. Within this AOI group the abstracted added design received more mean gaze time than the control. The number of document transitions, time on task, and relevancy responses, helped show how the design change affected undergraduates document triage behavior and their perception of relevancy. Of these data the only difference between the control and experimental groups was a reduced mean of surrogate stage transitions in the experimental environment. The survey data reported the demographics of the participant population sampled and their subjective experiences of the two different SERP design environments. These responses showed a diverse sampling of undergraduates at varying levels of research experience and most indicated a moderate familiarity with the tested topic, cyberbullying. Many participants had a high number of credit hours prior to the University of Missouri fall semester and were older freshmen. The survey response also showed that most participants preferred the experimental abstract design, especially if they had to use GS again for an assignment. As a total group most inaccurately predicted that they were faster in one design or another, at the individual level, only about 40% were able to correctly estimate which design was faster for them. The perception of improved confidence was also highly inaccurate, with only 25% of those who felt more confident in the shorter meta descriptions actually performing better in that design. A few more, 33%, of those of those who felt more confident in the longer meta
descriptions performed better in that design. The results also showed that perceptions of how often documents were opened were more accurate for the group that felt they opened more in the shorter description, 65%, than those who reported opening more with the longer descriptions, 40%. 
CHAPTER 5: DISCUSSION

Much previous research has found that freshmen undergraduates have been identified as a group who struggle with making sense of search results when doing research for class assignments, as explained in this dissertation’s introduction. This study replaced the existing Google Scholar (GS) search engine results page (SERP) meta descriptions with the abstracts of the corresponding retrieved articles and tested the effects of the SERP design change with an experiment. This chapter discusses the findings:

• RQ1: How does augmenting the meta description with an abstract in a SERP affect document triage behavior in undergraduates? Section 5.2 addresses the findings relating to the first research question using the group area of interest (AOI) data, document transition data, the total time on task, relevancy scores, and relevant survey data.

• RQ2: How does augmenting the meta description with an abstract in a SERP affect document triage behavior in undergraduates at the surrogate triage stage? Section 5.3 discusses findings and answers the second research question using AOI gaze data, document transition data, and the relevant survey data.

• RQ3: How does augmenting the meta description with abstracts in a SERP affect perception of document relevancy? Section 5.4 discusses how the relevancy scores and related survey results answer the third research question and relevancy perceptions.

• RQ4: How does augmenting the meta description with abstracts in a SERP affect document triage efficiency? In section 5.5 the fourth research question is
answered with a discussion of the time on task, relevancy scores, and relevant survey data.

This chapter ends with an exploration of future research possibilities and chapter conclusions.

**Undergraduate Freshmen Document Triage Behavior**

Of the research questions RQ1 was the broadest and most general of the four. In order to answer this question fully several different types of data collected in the study are used. Throughout the rest of this section each type of data and the hypothesis tested by that data are discussed in turn. First the AOI data from each AOI group is discussed followed by a discussion of the document transitions. Next the results involving the total time on task finding are discussed, followed by discussions of relevancy scores, and relevant survey data. The first type of data used to answer this question is the gaze time in the areas of interest. When all these contributing types of data have been analyzed the research question is answered.

**AOI Data**

The results show there are no relevant differences in the amount of gaze time per visit in the ‘title’ AOI group in the control and experimental designs. The results also show that there are no relevant differences in the amount of gaze time spent per visit in ‘author/publisher’ group AOI information in the control and experimental designs. The replacement of the normal ‘meta description’ with abstracts did significantly increase the amount of mean gaze time per visit for the AOI group in the experimental design. The mean gaze time per visit to the control meta description were 27.98 seconds, while in the experimental group this time per visit extended to 1:44.62. The longer descriptions were viewed and participants spent more time looking at them. This was not the case for the last ‘cited’ AOI group. The results show that there were no relevant differences in the amount of gaze time per visit in the ‘cited’ AOI group in the control and
experimental designs. Two specific hypotheses are tested with these AOI data. The first tests the designs effect on gaze time in the title AOI and the second tests the design change effect on the gaze time in the meta description AOI.

**Hypothesis 1.** $H_0$: There is no significant relationship between the inclusion of abstracts within Google Scholar SERP and the amount of time titles are viewed. Due to the data collected this null hypothesis has to be accepted; there was no significant change in how long the titles were viewed in the two SERP designs. Considering the prior research (Lozides & Buchana, 2001; Saracevic, 1969, Cool et al. 193) showed that the title was one of the most viewed elements of an academic paper while making relevancy decisions, by having a longer meta description on the SERP the title could have received more gaze time if SERP users did not need to investigate an article further by opening other document surrogates, like vendor pages or full text documents, but that was not the case.

**Hypothesis 2.** $H_0$: There is no significant relationship between the inclusion of abstracts within Google Scholar SERP and the amount of time meta descriptions are viewed. As stated above, there was a significant difference between the mean gaze time per visit in the group AOI meta description in the SERP control and SERP experimental conditions, with the experimental design receiving more time. Therefore the null must be rejected and we have to accept the alternative hypothesis, $H_1$: There is a significant relationship between the inclusion of abstracts within Google Scholar SERP and the amount of time meta descriptions are viewed.

These data bring new knowledge in that we now have some preliminary data about how long freshmen spend looking at the title, authors, meta descriptions and citation information elements on an academic SERP. The most important finding shows that the participants did change how long they looked at the expanded meta description. While there is more to read in a
longer meta description, it was unknown if participants would actually spend more time reading the description, especially when they are accustomed to having to go to another page to see a full abstract. This supports document triage findings that abstracts are a key component used while determining relevancy (Lozides & Buchana, 2001; Saracevic, 1969, Cool et al. 1993). These findings also support the work Cutrell and Guan did in 2010 which gathered eye tracking data on title, meta description, and URL components for MSN searches. They also found that longer meta descriptions increased gaze time spent within the meta description AOI.

Even though the gaze data for the author/publisher and cited elements were similar for both designs, the results show that these areas are viewed by freshmen rather than ignored or simply gazed at for little time as their view passed through one element to get to another. The results show 8.63 seconds mean time in the control design and 7.8 seconds mean time in experimental design within the author/publisher element. This is also true for the amount of time spent in the cited element, which recorded a mean time of 5.19 seconds mean time in the control design and 3.87 seconds mean time in the experimental design. It is important to note that gaze time alone does not tell us if these sections were read for meaning or if they were a part of the cognitive decision making process.

These findings also support research showing that as meta description get longer the amount of time spent looking at other elements gets shorter (Cutrell & Guan, 2010). The mean gaze time in the title element was reduced from 20.40 seconds in the control design to the 15.36 in the experimental design. The same was true for the author/publisher, which was reduced from a mean time of 08.63 seconds in the control design to 07.80 seconds in the experimental design. The cited element saw a reduction of mean gaze time from 05.19 seconds in the control design to 03.87 in the experimental design. From these data and the data above we can see that the design
change did affect gaze behavior in that it significantly increased gaze time in the meta-
description element while noticeably reducing it in others (see figure 5.1).

Figure 5.1 Mean Time per AOI Visit

**Document Transitions**

Each of the eye tracking sessions were reviewed in order to find and record all of the
transitions between documents surrogates in both the control and experimental designs in order
to answer RQ1. The results show that at surrogate level of triage there was a significant
difference in the mean number of transitions between the SERP control (M=6.84) and SERP
(M=4.81) experimental designs, with fewer surrogate transitions taking place in the experimental
design. This same observational data was also used to determine the number of transitions from
the SERP or surrogate documents to full text documents. These results show that there was no
significant difference in the mean number of transitions to the ‘within document stage’ between
the SERP control (M=.84) and SERP (M=.65) experimental designs. Two specific hypotheses
were tested with these document transition data. The first tests the design’s effect on the number
of ‘surrogate document’ stage transitions and the second tests the design change effect on the number of transitions to the ‘within document triage’ stage.

Hypothesis 3. H₀: There is no significant relationship between the inclusion of abstracts within Google Scholar and the median number of transitions between surrogate triage stage documents. Based on the results, the null must be rejected and the alternative hypothesis must be accepted. H₁: There is a significant relationship between the inclusion of abstracts within GS and the median number of transitions between surrogate triage stage documents. In Loizides’ Funnel model (2012) the surrogate triage stage is when an information seeker is encountering document surrogates, like SERPs or academic article vendor pages, which do not display full documents themselves. While no other known study tested document transitions based on SERP design in this way, these findings support Bae et al’s (2005) work showing that display configurations have an effect on document triage transitions. The Bae et al (2005) study investigated how multiple display types and availability affected document transitions, finding that when using a multi-display configuration the numbers of transitions increase over the number of transitions made using a single display. This dissertation shows new insight that the display design affects the number of documents transitions as well, showing that lengthier meta descriptions decreases the number of transitions at the surrogate level for most SERP users. A particular type of participant does not match this norm. While the majority of the participants did make fewer transitions at this document triage stage when using the longer meta descriptions, 8 participants, or 25%, use what I will call a completionist strategy and opened every link to the vendor surrogate via each article title regardless of what design they were using.

Hypothesis 4. H₀: There is no significant relationship between the inclusion of abstracts within Google Scholar and the median number of transitions between within document triage
stage documents. As the results show this hypothesis must be accepted in the null due to there being no significant relationship between the inclusion of abstracts within Google Scholar and the median number of transitions between within document triage stage documents.

The observational data also reveals some trends about how these full text documents were accessed. In total amongst all the participants in both control groups and experimental groups, 47 full text documents were viewed. This transition to a full text document is considered a separate stage of the document triage process in Loizides’ funnel model (2012) called the ‘within document stage’. Of these 47 full text document views only 4, 9%, were directly accessed by the links to full text documents provided on the GS SERP. As the heat map data showed, this part of the SERP was seen, just not used (see figure 5.2).

![Control](image1.png) ![Experimental](image2.png)

Figure 5.2 Group 1 Control and Experimental Heat Maps

A few, 6, 12%, were accessed from links within the document surrogate pages provided by the articles’ vendors. The remainder, 78%, of the full text document views came from the articles
whose vendors linked directly to a page, that along with all of the vendors’ Web designs elements, provided the document’s full text, rather than just the abstract like most of the vendors. If the vendor page was accessed via the SERP title element and had the full text, users were much more likely to scroll down and view elements beyond the title and abstract provided by most vendors. These were only considered to be links to full text views if the eye tracking showed participant fixations below the abstract; if the eye tracker did not show fixations below the abstract then the views of these surrogate full text hybrid pages were scored as surrogate transitions instead.

This observational data supports previous findings that areas like the title and abstract receive the most attention (Loizides & Buchanan, 2011; Saracevic, 1969, Cool et al. 1993) because most participants made relevancy decisions with only the information provided on the SERP or the document surrogates pages that only had title, unaltered meta description, and abstract information. The reduced number of full text document triage transition could also have been affected by the study being conducted on a laptop with only one screen, which other document triage research has shown to reduce the number of document triage transitions (Bae et al, 2005). These findings also support the study that showed that convenience is a critical factor in information-seeking behaviors (Connaway, Dickey & Radford, 2011), considering a click was too much effort to view a full text document for most of the participants but if the full text was only a scroll or glance away participants were more likely to take that step. The addition of abstracts to the SERP did change user behavior by reducing how often they felt the need to go seeking further information beyond what the SERP provided, but did not change their behavior relating to how often they sought out full text documents. Both of these findings support the principle of least effort (Zipf, 1949).
Total Time on Task

The mean total time on task scores are another data set that contribute to RQ1. The results showed that the design did not have a significant effect on the mean total time on task, with the mean times in the control (M=7:37.88) and experimental groups being (M=7:12.54) virtually the same. These data test a hypothesis relating to how the design change affected the time it took participants to complete the document triage task.

Hypothesis 5. H₀: There is no significant relationship between the inclusion of abstracts within Google Scholar and the total amount of time required to complete the research task. Due to the results the null has to be accepted: there was no significant relationship between the inclusion of abstracts within Google Scholar and the total amount of time required to complete the research task. These findings conflict with the work of Cutrell and Guan (2007) which found that longer meta description showed a reduction in task time for informational tasks. Informational tasks used in many SERP studies (Dumais, Buscher, & Cutrell, 2010, Cutrell and Guan 2007; Pan et al., 2007; Joachims, Granka, & Pan, 2005) asked participants to answer specific questions, e.g. ‘where is the tallest mountain in Missouri located?’ and not make decisions about a document’s relevancy relating to a topic. These findings do support a document triage research study, which used a relevancy sorting task like the one used in this experiment, that found people tend to make decisions about topic relevancy in less than one minute (Buchanan & Loizides, 2007). The abstract version of GS did not contribute to any sort of change in time on task for relevancy decision-making. These findings are discussed more thoroughly below in the section 5.5

Relevancy Scores
To satisfy RQ1 the data concerning the altered meta descriptions and relevancy scores are taken into account. These results showed no significant difference between the SERP control (M=83.44, SD=15.783) and SERP experimental conditions (M=82.81, SD=11.705). These data tested a final hypothesis.

**Hypothesis 6.** H₀: There is no significant relationship between the inclusion of abstracts within Google Scholar and the scores for relevancy accuracy. Because of these results this hypothesis must be accepted. This also conflicts with the work of Cutrell and Guan (2007), which found that with the use of a longer meta description participants showed an improvement in scores for informational tasks. Due to these results it can be said that the abstracts on the GS SERP did not change how accurately they sorted the documents into relevancy categories. More will be discussed about these findings in section 5.5.

**Relevant Survey Questions**

The final set of data used to answer RQ1 are the survey questions which asked if a participant could perceive a difference in the time it took to complete the task and discern how many document transition they made. These data are also important in that it has been shown that user preference and performance do not always correlate, and users can choose design over performance (Neilsen & Levy 1994). These data help show if their users were able to note any performance changes.

According to the survey results as a total group, most inaccurately predicted that they were faster in one design or another when the mean times are taken into account. At the individual level, only about 40% from either group were able to correctly estimate which design was faster for them. From these data it seems that the perception of the passage of time while performing document triage tasks is inaccurate for most freshmen.
The survey results also showed that perceptions of how often documents were opened were more accurate for the group that felt they opened more in the shorter description, 65%, than those who reported opening more with the longer descriptions, 40%. It seems like it was easier for people to notice a change in their own behavior when comparing how often they made document transitions if they self-reported opening more when in the design with shorter meta descriptions. With these two groups combined, the number of those who could accurately discern with which design they were making more document transitions was 59%. From these data it seems like it is easier for freshmen to discern how often they make document transitions than it is for them to gauge how long it takes them to make relevancy decisions.

**Answers to RQ1**

Now that all the different types of data used to answer RQ1 have been discussed, we can see that augmenting the meta description with an abstract in a SERP affects document triage behavior in undergraduates in several ways. The AOI data shows that the participants did change how long they looked at the expanded meta description. The data also showed another affect in gaze behavior in that it significantly increased gaze time in the meta description element while noticeably reducing it in others. The addition of abstracts to the SERP changed user behavior by reducing how often they felt the need to go seeking further information beyond what the altered SERP provided. These data indicate that the freshmen sampled more accurately discerned how often they make document transitions than they were able to gage how long it took them to make relevancy decisions. But while the design change affected document triage, it doesn’t affect behavior in several other ways.

The design change did not change participant behavior relating to how often they sought out full text documents or affect the time it took them to complete the document triage task. The
abstract version of GS did not contribute to any sort of change in time on task for relevancy decision-making, nor did their accuracy of sorting the documents into relevancy categories change.

**Undergraduate Freshmen Document Surrogate Triage Stage Behavior**

RQ2 asked: How does augmenting the meta description with an abstract in a SERP affect document triage behavior in undergraduates at the surrogate triage stage? This question was answered with the following data elements, the Group AOI findings, the number of document transitions, and the participants’ perceptions of how often they made document transitions.

As discussed above, the biggest changes in participant’s behavior relates to which elements of the SERP design attracted attention. The first $H_0$ relating to the mean gaze time at title element is accepted, while the second $H_0$ relating to the effect on mean gaze time in the amount of time meta descriptions are viewed GS is rejected. The design change did affect participant gaze behavior in that it significantly increased their gaze time in the meta description element, while noticeably reducing it in all of the other group AOIs.

Also the addition of abstracts to the SERP did change user behavior at the surrogate triage level, limiting their document transitions at this stage. This required the rejection of the third $H_0$: There is no significant relationship between the inclusion of abstracts within Google Scholar and the median number of transitions between surrogate triage stage documents. When asked about their perception of how many transitions they could remember making in each of the two SERP designs, most of the participants noted this change. 59% of surveyed participants were able to accurately access that they were making fewer transitions when using the abstract meta-descriptions. The results also showed that perceptions of how often documents were opened were
more accurate for the group that felt they opened more in the shorter description, 65%, than those who reported opening more with the longer descriptions, 40%.

The observational data also showed that users rarely left this stage of the triage process. Only 10 out of 47 transitions to the within document stage were initiated by click through to a full text document. The other 37 full text document views were on the vendor pages that had the full text already loaded into the part of their page that usually only contains the title and abstract. The data showing that without the convenience of these articles being a downward glance and a scroll away, it is less likely that they would have been viewed, support Connaway, Dickey & Radford’s (2011) findings of convenience being a key factor in research and use of resources. These findings also support Bae et al’s (2005) work showing that display configurations have an effect on document triage transitions giving us new understanding that the display design affects the number of documents transitions as well. As discussed above, a particular type of participant does not match this norm. 8 participants, or 25%, use what I will call a completionist strategy and opened every link to the vendor surrogate via each article title, regardless of what design they were using.

Taking these findings into account we can answerer RQ2 by saying: Augmenting the meta description with an abstract in a SERP affects document triage behavior in undergraduates at the surrogate triage stage by reducing the number of transitions they make to other surrogate documents, unless they are a particular type of user who opens every article title link to the next surrogate document level, the vendor article page. This reduction in surrogate stage transitions was noticeable by more than half of the participants. The design change did not change how often users left the surrogate stage and transitioned into the full text stage.

Undergraduate Freshmen Relevancy Perceptions
RQ3 asked: How does augmenting the meta description with abstracts in a SERP affect perception of document relevancy? This question is answered using relevancy scores and related survey results. As discussed above the data show no significant change in the relevancy sorting task scores, which forces the acceptance of the sixth $H_0$: There is no significant relationship between the inclusion of abstracts within Google Scholar and the scores for relevancy accuracy. This conflicts with Cutrell and Guan’s (2007) SERP study which found that a longer meta description produced a reduction in task time for informational tasks, but this conflict could relate to the differences in tasks. Informational tasks are used in many SERP studies (Dumais, Buscher, & Cutrell, 2010, Cutrell and Guan 2007; Pan et al., 2007; Joachims, Granka, & Pan, 2005), asking participants to use SERPs to answer specific questions. This is different from a relevancy-sorting task, which asks participants to discover an overall meaning of an article’s topic. Document triage studies use relevancy-sorting tasks or ask participants to rank article relevancy and compare their results in either case to that of a subject matter expert. The difference in the nature of these two tasks, finding a specific piece of information and making general sense of an article’s topic relevancy, could be an explanation for why there was no change between the control and experimental design SERPs. Another reason for the lack of the longer descriptions effectiveness and mean scores of 8 out of 10 could be due to how the articles used in the SERPs were chosen for their lack of ambiguity. There were no articles in a third ‘not relevant’ category or with mixed findings that may have forced participants to make a best-fit decision. Perhaps increasing the task difficulty might have revealed some differences between the different designs in their possible effectiveness of supporting relevancy decisions. Considering the undergraduate freshmen participant pool, the task was deliberately kept relatively simple.
In view of how user perceptions and preferences can effect how useful participants find a design, three of the survey questions were also designed to help answer some basic questions about their experience completing the tasks (Nielson & Levy, 1994). The first asked, of the two versions of Google Scholar did you like one better than the other? The results indicate a general preference for the SERP design with abstracts, though a few preferred the shorter meta description design. When asked if they had to use GS for a research paper, the numbers changed slightly causing four participants to change from being undecided: two preferred the longer meta description, and two others preferred the shorter description. Considering that the design change did not have any statistical effect on relevancy decisions, these data are less critical in determining if user preferences conflicted with their performance.

The last survey question relating to this research question asked if the participants remembered feeling more confident with one version of Google Scholar than the other. The results show that while there was a majority preference for the abstract SERP design, it only bolstered confidence in about half of the participants. This confidence has little merit. Only 25% of those who felt more confident in the shorter meta descriptions actually performed better in that design, and only 33% of those who felt more confident in the longer meta descriptions had their confidence justified.

Taking all of these findings into account, RQ3 can be answered stating that while there is a preference for the abstract included SERP design, it did not affect participant’s ability to sort documents based on their topic relevancy, nor did it bolster confidence appropriately for most of the participants.

Undergraduate Freshmen Triage Efficiency
RQ4 asked: How does augmenting the meta description with abstracts in a SERP affect document triage efficiency? For the purposes of this study, efficiency was defined as the ratio of time to accuracy of the relevancy-sorting task. This research question is answered with a discussion of the time on task, relevancy scores and relevant survey data.

As the results and prior discussion showed, it can be said that the abstracts on the GS SERP did not change how accurately participants sorted the documents into relevancy categories. Nor did the abstract version of GS contribute to any sort of change in time on task for relevancy decision-making. This was not expected, considering the work of Cutrell and Guan (2007) found that longer meta descriptions corresponded with a reduction in task time and an improvement in accuracy for informational tasks. As stated above, these findings do support document triage research that found people tend to make decisions about topic relevancy in less than one minute (Buchanan & Loizides, 2007), which was the case in both conditions. As mentioned previously, this could have to do the difference in nature between discerning relevancy and an informational task. Also, the lack of efficiency improvement could have to do with the level of difficulty of this experiment’s relevancy task in that it did not improve participant’s relevancy decisions. In terms of time, it could be that relevancy decisions are made in less than a minute regardless of design features, considering Buchanan & Loizides (2007) study found this to be the case when comparing electronic and paper media.

The survey showed there was a majority preference for the abstract SERP design and if GS had to be used again for a writing assignment, the longer abstract design was again preferred by most. The two designs, while being equally efficient in relation to time, did seem faster to many in one design or the other, but when all the participants total times on task is taken into account, these perceptions were illusionary. As a total group most, 40% accurately predicted that
they were faster in one design or another; at the individual level, only 40%, were able to correctly estimate which design was faster for them regardless of which design they felt was faster.

While not originally considered data to be used in answering RQ4, the addition of abstracts to the SERP changed user behavior by reducing how often they felt the need to go seeking further information beyond what the SERP provided. But while some might consider fewer clicks and page loads more efficient, the time savings were undetectable and not a part of this study’s efficiency definition.

These finding allows this RQ4 to be answered simply by stating, in no discernable way did augmenting the meta description with abstracts in a SERP change freshmen document triage efficiency.

**Future Research**

These findings do leave more questions. Considering this experiment and others (Buchanan & Loizides, 2007) have shown that document triage decisions tend to be made in less than a minute, can design improve document triage efficiency through reducing the time on task? Cutrell and Guan (2007) showed that longer meta descriptions improved both time on task and accurate resource selection for informational tasks. The research indicates there could be some sort of cognitive difference between the two tasks that allows for design change to improve things for one task but make no difference in the other. It could be revealing to study the differences, if any, in these two tasks cognitively. It could be interesting to run the same experiment and test informational task based questions rather than relevancy assessment questions to show more support for how these two tasks and SERP design effects are different.
Considering that it may not be possible to improve the efficiency of document triage in a
time on task way, it still may be possible that SERP design changes can affect the accuracy of
relevancy decisions. It would be worthwhile to try a similar experiment with a more complex
sorting task that was less scripted to see if the longer meta description had an effect, rather than
using this limited data set alone to support that claim. While all the data from this experiment are
drawn from one particular SERP, GS, it would be interesting to see if these mean scores differed
greatly in other academic SERP designs if they too had abstracts added to the experimental
groups meta descriptions.

It could also be of interest to compare undergraduates document triage performance with
that of graduate students, seeing just how much research experience has an effect. With the
addition of testing graduate students, it could be revealing to see how much the number of
citations effect relevancy with more complex relevancy tasks, through adjusting how often the
SERP shows a document has been cited. This would be similar to how several SERP studies
adjust rank order of results displayed to see what effect rank order has on tasks (Dumais,
Buscher, & Cutrell, 2010; Lorgio et al., 2008; Pan et al., 2007; Joachims et al., 2005). Citation
scores could reveal a similar bias. While the title and author did draw the most attention for
undergraduate freshmen, it would be worthwhile to see if the time in these same AOIs were
different for more graduate level researchers.

Another observation that needs further research is how different the heat map patterns
were in this study vs. other SERP studies. Often on SERP pages there is a heat map clustering
called the ‘golden triangle,’ which appears in the top left corner of the page centered on the first
result (Hotchkiss, Alston & Edwards, 2010). The heat maps presented here did not match this
pattern; was this due to the nature of the task or that the study asked for each result to be reviewed rather than inspected only if it caught the user’s interest?

The behavior that was most changed by augmenting the meta description with abstracts was the design reduced number of surrogate triage transitions. One particular group of participants showed no change in behavior and opened every article title link to the next surrogate document level regardless of which SERP design they were using. The experiments task was not naturalistic in that each resource had to be evaluated according to the task instructions. Does this completionist approach relate to a behavior pattern of a group of people who never make decisions based on SERP data alone and always look deeper, or was this a result of experiment’s task? Further research is required to discern if this is an actual type of information seeking behavior rather than a by-product of the experiment’s design.

Conclusions

This dissertation was guided by four research questions. RQ1 asks: How does augmenting the meta description with an abstract in a SERP affect document triage behavior in undergraduates? This dissertation’s findings revealed that the design change did affect user behavior in several ways. The AOI data show that the participants did change how long they looked at the expanded meta description. These AOI data also indicate another effect in gaze behavior in that it significantly increased gaze time in the meta description element, while noticeably reducing in the other three. The addition of abstracts to the SERP changed user behavior by reducing how often they left the SERP seeking further information when using the augmented design. The survey data indicate that more of the freshmen sampled accurately discerned how often they make document transitions than were able to make relevancy decisions.
RQ2 asked: How does augmenting the meta description with an abstract in a SERP affect document triage behavior in undergraduates at the surrogate triage stage? At the surrogate triage stage the design changed behavior by reducing the number of transitions participants make to other surrogate documents, unless they are a particular type of user this study labeled as completionists. This group of participant opens every article title link to the next surrogate document level regardless of which SERP design they are using. This reduction in surrogate stage transitions was noticeable by more than half of the participants.

RQ3 asks: How does augmenting the meta description with abstracts in a SERP affect perception of document relevancy? The relevancy scores and related survey results show that while there is a preference for the abstract included SERP design, it did not affect participant’s ability to sort documents based on their topic relevancy, nor did it accurately bolster confidence for most of the participants.

RQ4 was the final research question which asks: How does augmenting the meta description with abstracts in a SERP affect document triage efficiency? The time on task, relevancy scores, and relevant survey data allows this RQ4 to be answered simply by stating: augmenting the meta description with abstracts in a SERP change did not improve freshmen document triage efficiency.

All of these research questions were designed to help discover if the addition of abstracts to a SERP relating to academic articles could help improve undergraduate freshmen performance in making better sense of results while doing research, a task this group struggles with, as previous research has shown. Considering the answers to the four research questions we now know the augmentation of meta descriptions with article abstracts was preferred by a majority of
the freshmen undergraduates sampled, but while the design change did not show any signs of improving their document triage efficiency, it did not hurt their efficiency either.
CHAPTER 6: CONCLUSIONS

This dissertation presented an experiment using eye tracking techniques to investigate what effects a search engine results page design (SERP) change may have on freshmen undergraduate document triage behavior. Freshmen undergraduates have been identified as a group who struggle with making sense of search results when doing research for class assignments. Due to prior work showing that longer meta descriptions in SERPs improve people’s ability to answer information based questions and the importance of abstracts making relevancy decisions, this study augmented the existing Google Scholar (GS) SERP meta descriptions with the abstracts of the corresponding retrieved articles. Freshmen participants were tasked with reviewing ten documents relating to two topics involving cyberbullying in each design and sorting them into appropriate subject groups.

The study’s findings bring new understanding of how the design change affects gaze time in areas of interest (AOIs), showing that augmenting the SERP with an abstract caused the participants to look at the expanded meta description for longer times. The study found that the design change also affected gaze behavior in that it significantly increased gaze time in the meta description AOI while noticeably reducing it in the title, author and publisher information, and the number of citations and links to other works’ AOIs. This study brings new insight as to how the design change may affect undergraduates’ perception of which SERP results are relevant, and how it changes their document triage behavior.

The addition of abstracts to the SERP changed user behavior by reducing how often the sampled freshmen felt the need to go seeking further information beyond what the SERP provided, but did not change their behavior relating to how often they sought out full text documents. The abstract version of GS did not contribute to any sort of change in time on task
for relevancy decision-making. The study provided further evidence that convenience is a critical factor in information-seeking behaviors, due to convenience being the major factor contributing to when full text documents were accessed. Survey data suggest it was easier for people to notice a change in their own behavior when comparing how often they made document transitions, reporting opening more when in the design with shorter meta descriptions. The experiment also discovered a group of participants who used a completionist strategy, and opened every link to the vendor surrogate via each article title, regardless of what design they were using.

The design change did not have any effects on how long it took participants to complete their document triage tasks. According to the survey results as a total group, most participants inaccurately discerned that they were faster in one design or another. These findings further support other work showing that relevancy decisions are made in less than a minute (Buchanan & Loizides, 2007).

The experiment’s findings show that the addition of abstracts to the GS SERP did not change how accurately participants sorted the documents into relevancy categories. The survey results indicate a general preference for the SERP design with abstracts, though a few preferred the shorter meta description design. These data are also important in that it has been shown that user preference and performance do not always correlate, and users can choose design over performance (Neilsen & Levy 1994). In this case preference and performance are not a factor; regardless of which design was preferred the performance results were statistically the same.

For the purposes of this study, efficiency was defined as the ratio of time to accuracy of the relevancy-sorting task. Due to the design change showing no differences in the relevancy score or the total time on task, the change did not affect document triage efficiency. This study’s
findings conflict with Cutrell and Guan (2007), which found that longer meta descriptions corresponded with a reduction in task time and an improvement in accuracy for informational tasks. Due to this conflict, further research is needed to determine if informational tasks and relevancy tasks are so fundamentally different that one should not expect similar results, or if the document triage task was not challenging enough to show any benefits from longer meta descriptions.

Since the beginning of this project Summon, a federated database search service from ProQuest, has started to show the abstracts relating to retrieved items on its search results page. Its display is quite different from the SG experimental design tested in this dissertation, but due to this inclusion, the addition of abstracts to search results pages could be a coming trend making this research timely. Unfortunately we do not have access to the rational behind this change in summon service at this time. The advantages to having an abstract inclusive results page demonstrated in this study could be the reduction in click though to other pages, this reduces server loads by reducing server search requests.

But while this is a benefit to the system administrators to what extent does it help the user? This study’s findings showed that of those sampled the majority of users preferred the abstract inclusive design and subjective approval is not something to be dismissed from a usability stand point, nor is it easily dismissed by competitive search service providers be they web scale like GS or database orientated like Summon. The more people like a design the more likely they are to use it. The addition of abstracts lengthen the SERP page in the case of the GS design change. Could this be a problem? The research that has been done on the differences in reading on paging between documents and scrolling on the desktop and this body of work suggests that there is no statistically significant difference between the two page display
techniques for within document searching (Baker, 2003, Bernard, Baker and Fernandez, 2003, Eyuboglu and Orhan, 2011, Grace, 2005, Kim and Albers, 2001, Peytchev, Coupe, McCabe and Crawford, 2006 and Santosa, 2011). This supports including abstracts on results pages as they lack of click through to other pages and in increase in scrolling to view longer SERP displays is not negatively impacting researchers, though it might be anything to those with a personal preference for clicking rather than scrolling. Taking into account that some search products are already moving in the direction of including abstracts, that the abstract design is preferred, the additional scrolling for longer results pages is not harmful, and the benefits from a server administrator perspective, were I a SERP designer, even with what little evidence we have now I would think the addition of abstracts to results pages are beneficial.

These findings are likely to be of most interest to IR, document triage and HIB researchers. IR, SERP, and document triage researchers will be interested in how the design change affected click through behavior. Another point of interest to these researchers is the possible new type of triage behavior revealed by the data labeled completionist that has gone unnoticed up to this point. Both SERP and document triage IR researchers would be interested in how this study's findings further support earlier document triage research, providing further evidence that relevancy decisions tend to be made in less than a minute (Buchanan & Loizides, 2007). They would also take note of how the longer abstracts drew more attention but did not improve the simple sense making tasks’ results. As stated before, longer meta descriptions improved the accuracy of information tasks (Cutrell & Guan, 2007), but this was not the case for the simple sense making tasks used in this study. This would be of interest to both SERP and document
triate IR researchers, further demonstrating the relationship between design and specific
tasks and how a design improvement for one task may not necessarily aid in another.

HIB researchers are likely to be interested in how this study's findings further
support of the principle of least effort, as shown by the reduced click through behavior and
convenience being the key factor in why full text documents were opened. HIB researchers,
much like the SERP and document triage researchers, are also likely to be interested in that
this study showed continued evidence that relevancy decisions are typically being made in
less than a minute while performing document triage tasks. This data also supports the
principle of least effort. HIB researchers are also likely to take note of the potentially newly
discovered behavior pattern demonstrated by the group labeled completionists.

This study's findings are by no means conclusive and somewhat problematic,
considering the nature of the research task was scripted and kept purposely simple by only
using examples that were easily categorized. Further research is needed to discover if there
are benefits in more naturalistic uses of abstract enhanced search displays and with more
complex relevancy tasks. We might be able to say this study shows that with simple
relevancy sorting tasks the addition of abstracts has no effect in sense making ability, but
without repeating the study and attempting a similar study with larger sample group sizes
the overall evidence for even this claim is thin. Considering the advantage of being able to
compare multiple abstracts on the same page, one could make the intuitive leap that
abstracts on search results pages could in fact help with more complex search tasks. Only
more research will tell. Further research is also needed to test other academic SERP pages
like the one provided by Summon. Other participant groups, like graduate students or
faculty, need to be investigated to determine if research experience has an effect. With the
addition of testing faculty and graduate students with more complex relevancy tasks, specific AOI’s like citation elements need to be explored to see how much the number of citations effect relevancy, through increasing or decreasing how often the SERP shows a document has been cited.

The dissertation’s overall purpose was to test effects adding abstracts to a SERP relating to academic articles would have on undergraduate freshmen performance in making sense of results while doing research. Ultimately the design change did not show any evidence of improving freshmen’s document triage efficiency either in accuracy or total time on task, but it did not reduce their efficiency either. Considering the various findings we now know the augmentation of meta descriptions with article abstracts was preferred by a majority of the freshmen undergraduates sampled. The AOI data showed that the design change increased how long participants looked at the expanded meta description while noticeably reducing in the other three. The addition of abstracts to the SERP changed user behavior by reducing how often they left the SERP seeking further information when using the augmented design. The dissertations’ findings also suggest that a design change to vendor provided web pages ought to include full text of documents whenever possible rather than requiring additional steps. While the addition of abstracts did not help the sampled undergraduate freshmen with sense making with the study did provide other insights into design and its relationship to information seeking in this context.
REFERENCES


doi:10.1016/j.lisr.2011.02.001


APPENDIX A. IRB APPROVAL

March 19, 2015

Principal Investigator: Nathan John Lowrance
Department: Info Science & Learning Tech

Your Exempt Application to project entitled Google Scholar and Meta Descriptions: Does Adding Abstracts to Search Engine Page Results Aid in Undergraduate Document Triage Efficiency? was reviewed and approved by the MU Institutional Review Board according to terms and conditions described below:

IRB Project Number 2001848
IRB Review Number 202530
Approval Date of this Review March 19, 2015
IRB Expiration Date March 19, 2016
Level of Review Exempt
Project Status Active - Open to Enrollment
Exempt Categories 45 CFR 46.101b(2) Risk Level Minimal Risk

The principal investigator (PI) is responsible for all aspects and conduct of this study. The PI must comply with the following conditions of the approval:

1. No subjects may be involved in any study procedure prior to the IRB approval date or after the expiration date.
2. All unanticipated problems, adverse events, and deviations must be reported to the IRB within 5 days.
3. All changes must be IRB approved prior to implementation unless they are intended to reduce immediate risk.
4. All recruitment materials and methods must be approved by the IRB prior to being used.
5. The Annual Exempt Form must be submitted to the IRB for review and approval at least 30 days prior to the project expiration date. If the study is complete, the Completion/Withdrawal Form may be submitted in lieu of the Annual Exempt Form.
6. Maintain all research records for a period of seven years from the project completion date.
7. Utilize all approved research documents located within the attached files section of eCompliance. These documents are highlighted green.

If you have any questions, please contact the IRB at 573-882-3181 or irb@missouri.edu.

Thank you,

MU Institutional Review Board
APPENDIX B. SERP INSTRUMENTS

Group A Control:

cyberbullying prevention and effects

Scholar

About 10,000 results (0.88 sec)

Articles

Offline consequences of online victimization: School violence and delinquency

Psychology in the ... - Wiley Online Library

... A decade ago, technology had not advanced to the point where cyber bullying was even an issue ... fortunately, kids are keeping pace with the changes ... Report incident to cyber bully's parent ... Facilitate school climate encouraging students to report cyberbullying. 4 - Very ...

Cited by 270 Related articles All 1,200 versions Cite Save More

Tackling cyberbullying: Review of empirical evidence regarding successful responses by students, parents, and schools
S. Perren, L. Conron, M. Capp, P. Dubeur - International Journal of Information, Communication & Ethic ..., 2012 - ijic.org

... purposes. 2. Methodology Review of online literature on cyberbullying and internet harassment. The inclusion criteria for this meta-analysis were identified in relation to cyberbullying research on children ... Cyberbullying experiences, impacts and coping strategies as described by Australian Young ...

Cited by 31 Related articles All 1,200 versions Cite Save More

The role of school psychologists in the assessment, prevention, and intervention of cyberbullying
T. Diamantoulis, E. Donnou - Psychology in the ... - Wiley Online Library

... psychologists can collaborate with teachers to design lessons around the topic of cyber-bullying. Cyberbullying is equally imperative that the policy also address cyberbullying activities occurring off...

Cited by 100 Related articles All 1,700 versions Cite Save More

The emotional impact on victims of traditional bullying and cyberbullying: A study of Spanish adolescents

... reported being victims of both types of bullying (traditional and cyberbullying) and ... In order to examine cyberbullying, two types of cyberbullying were distinguished: aggression using of the emotional effects caused by indirect bullying (traditional and cyberbullying) in contrast to ...

Cited by 121 Related articles All 1,200 versions Cite Save More

Under the radar: Educators and cyber bullying in schools

... bullying to count a single incident as an experience of cyber bullying. 370 ... redirecting feelings; or external - no consequences, non-confrontational. When cyberbullying did not occur, many intervention or prevention programs exist that deal specifically with cyberbullying; a Campbell ...

Cited by 31 Related articles All 2 versions Cite Save More

The nature of cyberbullying, and strategies for prevention

... bullying to count a single incident as an experience of cyber bullying. 370 ... redirecting feelings; or external - no consequences, non-confrontational. When cyberbullying did not occur, many intervention or prevention programs exist that deal specifically with cyberbullying; a Campbell ...

Cited by 71 Related articles All 9 versions Cite Save More

Associations among bullying, cyberbullying, and suicide in high school students

... findings suggest that some gender-specific strategies for prevention and intervention may be helpful. In contrast, we found that cyberbullying perpetration (not victimization) was a direct predictor of suicidal ideation. We speculate that for cyberbullying perpetration, a statement that was intended as a ...

Cited by 63 Related articles All 11 versions Cite Save More

Students' perspectives on cyber bullying

... phones on campus present this potential by ensuring that cellular phone policies are enforced with consistent consequences for students ... Finally, schools are encouraged to develop bullying prevention programs that include classroom lessons on cyber bullying to ensure ...

Cited by 276 Related articles All 15 versions Cite Save More
APPENDIX B. SERP INSTRUMENTS

Group B Control:

- Cyberbullying in schools a research of gender differences
- Cyberbullying, school bullying, and psychological distress: A regional census of high school students
- Cyberbullying: What school administrators (and parents) can do
- Cyberbullying prevention program
- Psychological, physical, and academic correlates of cyberbullying and traditional bullying
- Following you home from school: A critical review and synthesis of research on cyberbullying victimization
- A useful evaluation design, and effects of the Olweus Bullying Prevention Program

[Search results for related terms]
APPENDIX B. SERP INSTRUMENTS

Group B Control (continued):

Is cyberbullying worse than traditional bullying? Examining the differential roles of medium, publicity, and anonymity for the perceived severity of bullying
F Skoda, G Pomer - Journal of youth and adolescence, 2013 - Springer
...Therefore, awareness about the nonexistence of anonymity in cyberspace pays a double role in the prevention of cyberbullying and is also very important in traditional bullying. ... Cyber bullying
Bullying in the digital age ... Research into cyberbullying: Context ...
APPENDIX B. SERP INSTRUMENTS

Group A Experimental:

```
<table>
<thead>
<tr>
<th>Scholar</th>
<th>About 10,000 results (0.88 sec)</th>
<th>My Citations</th>
</tr>
</thead>
</table>
| Articles | Offline consequences of online victimization: School violence and delinquency S. Middle, D. Parkinson - Journal of school violence, 2007 - Taylor & Francis
As increasing numbers of youth embrace computer-mediated communication to meet academic and social needs, interpersonal violence directed and indirectly related to the Internet is occurring more often. Cyberbullying in particular has shot to the forefront of agendas in schools and communities due to the emotional, psychological, and even physical harm to which victims can be subjected. While previous studies have focused on describing its frequency in an exploratory capacity, the current work seeks to utilize general social theory to identify the emotional and behavioral effects of cyberbullying victimization. Data collected online from a sample of adolescent Internet-users indicate that cyberbullying is a potent form of strain that may be related to involvement in school problems and delinquent behavior offline. Implications of these findings and suggestions for policy are discussed.
Cited by 280 Related articles All 6 versions Cite Save More

PDF by byu.edu
Find in MU

High school teachers' perceptions of cyberbullying prevention and intervention strategies S. Stadler, M. Hoehn, D. Owen - Psychology in the ... - Wiley Online Library
Recent meta-analyses indicate that bullying prevention programs produce minimal change in student behavior. This study examined 60 high school teachers' perceptions regarding the effect of cyberbullying on students, which is a growing problem in the current digital age. The study found that teachers believed that cyberbullying prevention programs were effective in reducing instances of bullying. Teachers also believed that prevention programs should include strategies such as identifying perpetrators, implementing school policies, and providing support services. The study highlights the importance of teacher perception in the effectiveness of bullying prevention programs.
Cited by 27 Related articles All 4 versions Cite Save More

div class="qss_gpl_F">

Tracking cyberbullying: Review of empirical evidence regarding successful responses by students, parents, and schools S. Parnes, L. Conron, H. Owen, F. Denne - International Journal of ... - iop.org
A summary of current knowledge on the successful responses to cyberbullying. The study found that successful responses to cyberbullying include reducing risks, combating the problem, and buffering negative impact. A systematic literature review of 30 studies, most of which report findings regarding general prevention strategies (e.g., anti-bullying policies or school safety programs) and the use of coping strategies such as seeking support, responding (negotiation or confrontation), technical solutions, and avoiding and emotion-focused strategies. While a few studies report perceived success, few measure the success of these strategies in relation to risks and outcomes. There is a clear lack of evidence concerning successful responses.
Cited by 31 Related articles All 12 versions Cite Save More

The role of school psychologists in the assessment, prevention, and intervention of cyberbullying T. Diamandouros, D. Horner - Psychology in the ... - Wiley Online Library
Many school-aged children experience the reality of bullying. The impact and effects of bullying have been described extensively in scholarly literature. Unfortunately, the impact of bullying has also made contemporary headlines with the recent rash of school shootings and other forms of school-based violence. The current body of literature brings an additional threat to bullying. Due to the rapid pace of technological advances, today's students are using contemporary delivery systems to taunt, tease, and threaten their classmates. Cyberbullying is the newest form of bullying, and it brings additional challenges for school personnel. Because of the unique aspects of cyberbullying, school psychologists need strategies to guide school communities faced with this threatening behavior.
Cited by 110 Related articles All 7 versions Cite Save More

We examined the emotional impact caused to victims of bullying in its traditional form, both directly and indirectly, as well as bullying inflicted by use of new technologies such as mobile phones and the Internet. A sample of 1,671 adolescents and young people responded to a questionnaire when asked to report if they had been victims of various forms of bullying, as well as the emotions this caused. The results show that although traditional bullying affected significantly more young people than cyberbullying, the latter affected one in ten adolescents. Analysis of the emotions caused showed that traditional bullying produced a wider variety of impacts, with the victims being divided into five different emotional categories, while indirect bullying and cyberbullying presented a narrower variety of results with the victims being classified into just two groups. Those who said that they had not been emotionally affected and those who simultaneously suffered both forms of bullying.
Cited by 21 Related articles All 4 versions Cite Save More
```

117
APPENDIX B. SERP INSTRUMENTS

Group A Experimental (continued):

'Under the radar': Educators and cyberbullying in schools

Educators play an important role in preventing cyberbullying and encouraging positive online behaviour. In this study we report on the educator portion of a study of cyberbullying at five large, technology-rich secondary schools in Canada. We discuss 17 educators' experiences with cyberbullying, their knowledge of social networking technology, their awareness of the extent of cyberbullying among their students, and their role in detecting and preventing cyberbullying. We also discuss qualitative analyses of teachers' responses to open-ended questions. The data revealed that educators are aware of the extent of cyberbullying among their students and that they see prevention as a priority, and were able to pose possible solutions, policies or strategies that could be implemented. This study was carried out in the last decade and was illuminated by the younger teachers, who were more technologically savvy. Not all were educators interested in learning the results of their student portion of our research, preferring instead that educators remain under their radar.

Cited by 31 Related articles All 2 versions Cite Save

The nature of cyberbullying, and strategies for prevention
R Gown, M G Smith - Computers in Human Behavior, 2013 -Elsevier

Cyberbullying has been considered an important issue in the last decade. This paper reviews some recent findings and discusses general concepts within the area. The review covers definitional issues such as repetition and power imbalance, types of cyberbullying, age and gender differences, overlap with traditional bullying, and sequence of events, differences between cyberbullying and traditional bullying, motives for and impact of cyber victimization, coping strategies, and prevention intervention possibilities. These issues will be illustrated by reference to recent and current literature, and by in-depth interviews with nine Swedish students aged 13–15 years, who had some first-hand experience of one or more cyberbullying episodes. We conclude by discussing the evidence for different coping, intervention and prevention strategies.

Cited by 71 Related articles All 8 versions Cite Save

Associations among bullying, cyberbullying, and suicide in high school students

This study examined associations among depression, suicidal behaviors, and bullying and victimization experiences in 1491 high school students using data from the 2006 Youth Risk Behavior Survey. Results demonstrated that depression mediated the association between bullying/victimization and suicide attempts, but differently for males and females. Specifically, depression mediated the link between traditional victimization and suicide attempts similarly across gender, whereas depression mediated the link between cyber victimization and suicide attempts only for females. Similarly, depression mediated the link between traditional bullying and suicide attempts for females only. Depression did not mediate the link between cyberbullying and suicide attempts for either gender. Implications of the findings are discussed, including the importance of greater detection of depression among students involved in bullying, and the need for suicide prevention and intervention component in bullying programs. Findings suggest that bullying prevention efforts should be expanded from middle school students to include high school students.

Cited by 53 Related articles All 11 versions Cite Save

Students' perspectives on cyber bullying

The aim of this study was to gain a better understanding of the impact of cyber-bullying on students and the possible reasons for preventive measures targeting students, educators, and parents. A total of 148 middle and high school students were interviewed during focus groups held at two middle and two high schools in a public school district. The focus groups were approximately 45 minutes in length. Students were divided by gender and asked a series of scripted questions by a same-gender student assistance counselor. We found that students' comments during the focus groups suggest that students—particularly females—view cyber bullying as a problem, but one rarely discussed at school, and one that should be stopped. The school district personnel as helpful resources when dealing with cyber bullying. Students are currently experiencing the majority of cyber bullying instances outside of the school day, however there is some impact at school. Students were able to suggest some basic strategies for dealing with cyber bullying, but were less likely to be aware of strategies to request the removal of objectionable websites, as well as how to respond as a helpful bystander when witnessing such online behavior. We conclude that school districts should address cyber bullying through a combination of policies and information that are shared with students and parents. Schools should include cyber bullying as part of their bullying prevention strategies and include classroom lessons that address reporting and bystander behavior.

Cited by 276 Related articles All 15 versions Cite Save

Cyberbullying Prevention: One Primary School’s Approach,

Bullying in all its forms, including cyberbullying, is a continuing problem in schools. Given the severe consequences it can have on students (socially, psychologically and physically) it is not surprising that a number of intervention programs have been developed, with most adopting a whole-school approach. The current study compared students' self-reports on bullying between schools with and without a Prevention of Cyberbullying (PCB) approach. A sample of 36 students in the PCB school and a matched sample of 35 students in other schools between the ages of 10 and 13 completed the Student Bullying Survey. Results indicated that while there were significant differences in incidence of face-to-face bullying, there were similar results from both cohorts in relation to cyberbullying. Both groups of students felt that teachers were more likely to prevent face-to-face bullying than cyberbullying. Findings indicate that teachers and guidance counselors need to be involved in teaching strategies about cyberbullying as they are in teaching strategies about reducing face-to-face bullying. Findings indicate that teachers and guidance counselors need to be involved in teaching strategies about cyberbullying as they are in teaching strategies about reducing face-to-face bullying.

Cited by 16 Related articles All 8 versions Cite Save
APPENDIX B. SERP INSTRUMENTS

Group B Experimental:

Cyberbullying prevention and effects

Scholar
Page 2 of about 10,900 results (0.88 sec)

<table>
<thead>
<tr>
<th>Articles</th>
<th>Cyberbullying in schools: a research in gender differences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The study investigates the nature and extent of adolescents' experience of cyberbullying. A survey study of 264 students from three junior high schools was conducted. In this article, cyberbullying refers to bullying via electronic communication tools. The results show that close to half of the students were bullied victims and about one in four had been cyber-bullied. Over half of the students reported that they knew someone being cyberbullied. Almost half of the cyberbullied used electronics means to harass others more than three times. The majority of the cyber-bullied victims and bystanders did not report the incidents to adults. When gender was considered, significant differences were detected in terms of bullying and cyberbullying. Males were more likely to be bullies and cyberbullying than their female counterparts. In addition, female cyberbullying victims were more likely to inform adults than their male counterparts.</td>
</tr>
<tr>
<td></td>
<td>Click for more</td>
</tr>
</tbody>
</table>

Cyberbullying, school bullying, and psychological distress: A regional census of high school students

<table>
<thead>
<tr>
<th>Articles</th>
<th>Cyberbullying, school bullying, and psychological distress: A regional census of high school students</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Objectives. Using data from a regional census of high school students, we have documented the prevalence of cyberbullying and school bullying victimization and their associations with psychological distress. Methods. In the fall of 2006, 25,406 ninth- through twelfth-grade students in Two-state Massachusetts completed surveys assessing their bullying victimization and psychological distress, including depressive symptoms, self-injury, and suicidality. Results. A total of 15.8% of students reported cyberbullying and 26.9% reported school bullying in the past 12 months. A majority (55.7%) of cyberbullying victims were also school bullying victims; 30.3% of school bullying victims were also cyberbullying victims. Victimization was higher among nonheterosexual-identified youth. Victims report lower school performance and school attainment. Controlled analyses indicated that distress was highest among victims of both cyberbullying and school bullying (adjusted odds ratios [AORs] were from 0.38 for depressive symptoms to 0.35 for suicide attempts requiring medical treatment). Victims of either form of bullying alone also reported elevated levels of distress. Conclusions. Our findings confirm the need for prevention efforts that address both forms of bullying and their relation to school performance and mental health.</td>
</tr>
<tr>
<td></td>
<td>Click for more</td>
</tr>
</tbody>
</table>

Defining cyberbullying: A qualitative research into the perceptions of youngsters

<table>
<thead>
<tr>
<th>Articles</th>
<th>Defining cyberbullying: A qualitative research into the perceptions of youngsters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Data from 63 focus groups, which involved students from 10 to 18 years old, show that youngsters often interpret cyberbullying as &quot;Internet bullying&quot; and associate the phenomenon with a wide range of practices. In order to be considered &quot;true&quot; cyberbullying, these practices must meet several criteria. They should be intended to harm (by the perpetrator) and perceived as harmful (by the victim), be part of a repetitive pattern of negative offline or online actions, and be performed in a relationship characterized by a power imbalance (based on &quot;real-life&quot; power criteria, such as physical strength or age, and/or on ICT-related criteria such as technological know-how and anonymity).</td>
</tr>
<tr>
<td></td>
<td>Click for more</td>
</tr>
</tbody>
</table>

Recommended practices: A review of schoolwide preventative programs and strategies on cyberbullying

<table>
<thead>
<tr>
<th>Articles</th>
<th>Recommended practices: A review of schoolwide preventative programs and strategies on cyberbullying</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cyberbullying is a growing concern among school-age students. The combination of increased access, ease, and use of Web-based communication are part of what makes this such a problem. Cyberbullying affects not only students, but also parents and administrators. Because schools have the ability to reach all groups involved, it may be practical for schools to implement preventive and supportive efforts to address this concern. Research shows that prevention programs focused at the school level will have the strongest effect. By implementing a consistent schoolwide approach to cyberbullying, it is hoped that the benefits will carry through a student's academic career and beyond. This article reviews recent research and Web-based resources on existing programs and makes recommendations for schools to develop their own cyberbullying prevention program.</td>
</tr>
<tr>
<td></td>
<td>Click for more</td>
</tr>
</tbody>
</table>

Cyberbullying: What school administrators (and parents) can do

<table>
<thead>
<tr>
<th>Articles</th>
<th>Cyberbullying: What school administrators (and parents) can do</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Technology has transformed the lives of adolescents, including the ways they bully one another. Various refer to the use of cell phones, instant messaging, Web sites, voicemail, and chat on social media to predict the impact on and parents. To combat cyberbullying, educators need to better understand the nature of it and be aware of actions that they can undertake to prevent cyberbullying in schools.</td>
</tr>
<tr>
<td></td>
<td>Click for more</td>
</tr>
</tbody>
</table>
APPENDIX B. SERP INSTRUMENTS

Group B Experimental (continued):

Knowing, building and living together on internet and social networks: The ConRod cyberbullying prevention program
R. Ottege-Ruc, R. De Hay, J. A. Geas - International Journal of Conflict and..., 2012 - Jour... An evaluation of the success of the evidence-based ConRod program, which addresses cyberbullying and other emerging problems linked with the use of the internet and seeks to promote a positive use of this new form of communication. The main aims of the ConRod program are a) to improve perceived control over information on the internet, b) to reduce the time dedicated to digital device usage, and c) to prevent and reduce cyberbullying. The impact of the program was evaluated with a quasi-experimental design with a sample of 923 students (556 experimental and 369 control). The results of the mixed repeated measures ANOVA's demonstrate that ConRod contributes to reducing cyberbullying and cyber-dependence, to adjusting the perception of information control, and to increasing the perception of safety at school.
Cited by: 11 Related articles | All 3 versions | Cite Save

(HM) Psychological, physical, and academic correlates of cyberbullying and traditional bullying
Purpose: To examine the relationship between children's and adolescents' experiences with cyberbullying and traditional bullying and psychological, physical, and academic performance. Methods: Nine hundred thirty-one students in grades 6 through 12 completed an anonymous survey examining their experiences with cyberbullying and traditional bullying. Also included were measures of anxiety, depression, self-esteem, physical wellbeing, school attendance, and academic performance. Results: Participants were categorized as belonging to one of four groups: cyber victim, cyber aggressor, bully-victims, and those not involved in cyberbullying. A similar categorization was done with traditional bullying. Those in the bully-victims group (and particularly the cyber-bullying victim group) had the most negative scores on most measures of psychological health, physical health, and academic performance. Conclusions: This appears to be somewhat, although not perfect, overlap between involvement in traditional bullying and cyberbullying. Additionally, the physical, psychological, and academic correlates of the two types of bullying remain one another.
Cited by: 66 Related articles | All 5 versions | Cite Save

Following you home from school: A critical review and synthesis of research on cyberbullying victimization
NS. Takakura - Computer in Human Behavior, 2010 - Elsevier
More than 10% of youths in the United States are connected to the Internet in some way. An unreported outcome of the Internet's pervasive reach is the growing rate of harmful offenses against children and teens. Cyberbullying victimization is one such offender that we recently received a lot of attention. The present report synthesizes findings from quantitative research on cyberbullying victimization. An integrative definition for the term cyberbullying is provided, differences between traditional bullying and cyberbullying are explained, areas of convergence and divergence are offered, and sampling and/or methodological explanations for the inconsistencies in the literature are considered. About 26–41% of all youths have experienced cyberbullying at least once in their lives. Demographic variables such as age and gender do not appear to predict cyberbullying victimization. Evidence suggests that victimization is associated with serious psychosocial, affective, and academic problems. The report concludes by outlining several areas of concern in cyberbullying research and discusses ways that future research can remedy them.
Cited by: 445 Related articles | All 4 versions | Cite Save

A useful evaluation design, and effects of the Olweus Bullying Prevention Program
D. Olweus - Psychology, Crime & Law, 2005 - Taylor & Francis
This design presents the logic and other characteristics of an "extended selection cohorts" quasi-experimental design. Possible threats to the validity of conclusions based on this kind of design are discussed. It is concluded that chances are good that conclusions about the effects or non-effects of school-based bullying programs will be roughly correct in most cases. The design may be particularly useful in studies where it is not possible or desirable to use a random selection of control schools and it should be used by both practitioners and researchers. The design is illustrated with a study in which three consecutive cohorts of students in approximately 21 (300 were administered the Bully/Victim Questionnaire before and after some 8 months of intervention with the Olweus Bullying Prevention Program (OBDP). Results indicated quite substantial reductions (24–49%) in bully/victim problems. The "time series" nature of the data strengthened confidence that a "history intervention" of the findings (Cook & Campbell, Quasi-experimentation, Chicago: Rand McNally, 1979) is very unlikely. The data in this project were obtained in the context of a government-funded "new national initiative against bullying in Norway. The characteristics of this initiative and the model used in implementing the program in more than 400 schools were briefly described.
Cited by: 129 Related articles | All 8 versions | Cite Save More

Is cyberbullying worse than traditional bullying? Examining the differential roles of medium, publicity, and anonymity for the perceived severity of bullying
H. Broek - Journal of Youth and Adolescence, 2019 - Springer
Cyberbullying: a modern form of bullying performed using electronic forms of contact (e.g., SMS, WMS, Facebook). It has been considered as being worse than traditional bullying in its consequences for the victim. This difference was mainly attributed to some specific aspects that are believed to distinguish cyberbullying from traditional bullying: an increased potential for a large audience, an increased potential for anonymous bullying, lower levels of direct feedback, decreased time and space limits, and lower levels of supervision. The present study investigated the relative importance of medium (traditional vs. cyber), publicity (public vs. private), and bullying anonymity (anonymous vs. not anonymous) for the perceived severity of hypothetical bullying scenarios among a sample of Swiss seventh- and eighth-grade students, mean age = 12.7, study 1: 46% female, mean age = 14.2). Participants ranked a set of hypothetical bullying scenarios from the most severe to the least severe one. The scenarios were experimentally manipulated based on the aspect of medium and publicity (study 1) and medium and anonymity (study 2). Results showed that public scenarios were perceived as worse than private ones, and that anonymous scenarios were perceived as worse than non-anonymous ones. Cyber scenarios generally perceived as worse than traditional ones, although effect sizes were found to be small. These results suggest that the role of medium is secondary to the role of publicity and anonymity when it comes to evaluating bullying severity. Therefore, cyberbullying is not a priori perceived as worse than traditional bullying. Implications of the results for cyberbullying prevention and intervention are discussed.
Cited by: 44 Related articles | All 14 versions | Cite Save
APPENDIX C. RELEVANCY TASK AND INSTRUMENT

Participant Instructions

For the purposes of this study you will be playing the role student with an assignment due in a psychology class.

Your instructor wants you to write a paper about both the effects of cyberbullying and what policies or prevention strategies may work.

Your job is to review 20 articles from Google Scholar about cyberbullying.

The search strings have already been created for you and 20 results have already been found. These 20 articles will be presented to you in two sets of 10 each.

For each set of 10 articles review and sort each of the retrieved articles into one of two categories that it fits best.

One category is if the article is mainly about how cyberbullying affects people.

The other category is if the article is mainly about cyberbullying prevention and policies.

Feel free to click links found on the page and use the interface as you would normally while making your decision about each result, but be sure to use the browsers back button to return to the search results screen.

On the list of printed article titles given to you will find that the articles are in the same order that they are on the screen and under each retrieved article are two choices, cyberbullying effects and cyberbullying prevention and policies.

For each article circle the one category that it best fits.

Remember the study is designed to test the interface, not you. Most people make relevancy decisions in less than a minute when looking for articles, so it is okay to go with your gut and quickly decide which articles are relevant or not.

Let the researcher know if you have any questions and if you are ready to begin.
APPENDIX C. RELEVANCY TASK AND INSTRUMENT

Offline consequences of online victimization: School violence and delinquency
Effects of cyberbullying Policies and Treatments

High school teachers' perceptions of cyberbullying prevention and intervention strategies
S Stauffer, MA Heath, SM Coyne… - Psychology in the …, 2012 - Wiley Online Library
Effects of cyberbullying Policies and Treatments

Tackling cyberbullying: Review of empirical evidence regarding successful responses by students, parents, and schools
S Perren, L Corcoran, H Cowie, F Dehue… - International Journal of …, 2012 - ijcv.org
Effects of cyberbullying Policies and Treatments

The role of school psychologists in the assessment, prevention, and intervention of cyberbullying
T Diamanduros, E Downs… - Psychology in the …, 2008 - Wiley Online Library
Effects of cyberbullying Policies and Treatments

The emotional impact on victims of traditional bullying and cyberbullying: A study of Spanish adolescents.
Effects of cyberbullying Policies and Treatments
APPENDIX C. RELEVANCY TASK AND INSTRUMENT

'Under the radar': Educators and cyberbullying in schools
W Cassidy, K Brown, M Jackson - School Psychology International, 2012 - spi.sagepub.com
  Effects of cyberbullying  Policies and Treatments

The nature of cyberbullying, and strategies for prevention
  Effects of cyberbullying  Policies and Treatments

Associations among bullying, cyberbullying, and suicide in high school students
  Effects of cyberbullying  Policies and Treatments

Students' perspectives on cyber bullying
  Effects of cyberbullying  Policies and Treatments

Cyberbullying Prevention: One Primary School's Approach
D Tangen, M Campbell - Australian Journal of Guidance and …, 2010 - Cambridge Univ Press
  Effects of cyberbullying  Policies and Treatments
APPENDIX C. RELEVANCY TASK AND INSTRUMENT

Cyberbullying in schools a research of gender differences
Q Li - School psychology international, 2006 - spi.sagepub.com

   Effects of cyberbullying       Policies and Treatments

Cyberbullying, school bullying, and psychological distress: A regional census of high school students
SK Schneider, L O'Donnell… - … Journal of Public …, 2012 - ajph.aphapublications.org

   Effects of cyberbullying       Policies and Treatments

Defining cyberbullying: A qualitative research into the perceptions of youngsters
H Vandebosch… - CyberPsychology & …, 2008 - online.liebertpub.com

   Effects of cyberbullying       Policies and Treatments

Recommended practices: A review of schoolwide preventative programs and strategies on cyberbullying
MA Couvillon, V Ilieva - Preventing School Failure: Alternative …, 2011 - Taylor & Francis

   Effects of cyberbullying       Policies and Treatments

Cyberbullying: What school administrators (and parents) can do

   Effects of cyberbullying       Policies and Treatments
APPENDIX C. RELEVANCY TASK AND INSTRUMENT

Knowing, building and living together on internet and social networks: The ConRed cyberbullying prevention program

R Ortega-Ruiz, R Del Rey, JA Casas - International Journal of Conflict and …, 2012 - ijcv.org

Effects of cyberbullying Policies and Treatments

Psychological, physical, and academic correlates of cyberbullying and traditional bullying


Effects of cyberbullying Policies and Treatments

Following you home from school: A critical review and synthesis of research on cyberbullying victimization

RS Tokunaga - Computers in Human Behavior, 2010 - Elsevier

Effects of cyberbullying Policies and Treatments

A useful evaluation design, and effects of the Olweus Bullying Prevention Program

D Olweus - Psychology, Crime & Law, 2005 - Taylor & Francis

Effects of cyberbullying Policies and Treatments

Is cyberbullying worse than traditional bullying? Examining the differential roles of medium, publicity, and anonymity for the perceived severity of bullying


Effects of cyberbullying Policies and Treatments
APPENDIX D. SURVEY QUESTIONS

1) Age
   1. 17-18   2. 18-19   3. 20-21   4. 21-older

2) Gender
   1. Male   2. Female   3. Transgender

3) Approximately (best estimate) how number of papers you have written for your college classes?
   1. 3-6   2. 7-9   3. 10-12   4. 13-15   5. 16 or more

4) How many people where in your high school graduating class?
   1. Less than 100
   2. 101-300
   3. 301-500
   4. 501-1000
   5. 1000 or more

5) Approximately (best estimate) how many credit hours have you completed, (not including this semester)?
   1. Less than 12   2. 12   3. 15   4. 18   5. 21 or more

6) How familiar are you with Google Scholar?
   1. I have never used Google Scholar before.
   2. I have used Google Scholar once or twice.
   3. I have used Google Scholar several times before.
   4. I use Google Scholar often.
   5. I almost always use Google Scholar.

126
APPENDIX D. SURVEY QUESTIONS

7) Did you already know allot about the topic of cyberbullying?

1. I have heard of it but I have never paid much attention to it.

2. I know a little about it, but I don’t know more than the basics.

3. I know some about it from the news and classes, enough that I could explain it to someone else.

4. I know more than most people from the news, classes, and my own interest in the topic.

5. I know allot about the topic, I have read articles, watched videos, or been to lectures about it.

The prior experiment had you place documents into two categories using two different version of Google Scholar. One had shorter descriptions and the results looked similar to this:

The other had longer descriptions and looked similar to this:
APPENDIX D. SURVEY QUESTIONS

The following questions have to do with the different versions of Google Scholar with longer descriptions and shorter descriptions and the sorting task.

8) Did you feel like you were faster in one version of Google Scholar or another?
   1. Neither was faster   2. The shorter descriptions   3. The longer descriptions

9) Of the two versions of Google Scholar did you like one better than the other?
   1. Neither.   2. The shorter descriptions.   3. The longer descriptions.

10) If you had to use Google Scholar for a class research paper would you prefer one version over the other?
    1. Neither.   2. The shorter descriptions.   3. The longer descriptions.

11) Did you remember feeling more confident with one version of Google Scholar than the other?
    1. I don’t remember feeling more confident with either version.
    2. I felt more confident with the shorter descriptions.
    3. I felt more confident with the longer descriptions.

12) Do you remember opening other webpages or PDF documents more often in one environment than the other?
    1. I don’t remember a difference with either version.
    2. I opened more with the shorter descriptions.
    3. I opened more with the longer descriptions.
A Spearman's rank-order correlation was run to determine the relationship between total time on tasks and the estimated number of papers written in the control condition at the p < .05 level of significance. There was no correlation between total time on tasks and the estimated number of papers written in the control condition, which was statistically significant ($r_s(8) = -.210, p = .266$).
A Spearman's rank-order correlation was run to determine the relationship between total time on tasks and the estimated number of papers written in the experimental condition at the $p < .05$ level of significance. There was no correlation between total time on tasks and the estimated number of papers written in the experimental condition, which was statistically significant ($r_s(8) = -.165, p = .385$).
A Spearman's rank-order correlation was run to determine the relationship between total time on tasks and graduating class size in the control condition at the $p < .05$ level of significance. There was no correlation between total time on tasks and the graduating class size in the control condition, which was statistically significant ($r_s(8) = -.243, p = .169$).
A Spearman's rank-order correlation was run to determine the relationship between total time on tasks and the estimated graduating class size in the experimental condition at the p < .05 level of significance. There was no correlation between total time on tasks and the estimated graduating class size in the experimental condition, which was statistically significant ($r_s(8) = .094, p = .623$).
A Spearman's rank-order correlation was run to determine the relationship between total time on tasks and number of credit hours taken in the control condition at the $p < .05$ level of significance. There was no correlation between total time on tasks and number of credit hours taken in the control condition, which was statistically significant ($r_s(8) = -.294, p = .115$).
A Spearman's rank-order correlation was run to determine the relationship between total time on tasks and number of credit hours taken in the experimental condition at the p < .05 level of significance. There was no correlation between total time on tasks and number of credit hours taken in the experimental condition, which was statistically significant ($r_s(8) = -.274$, $p = .143$).
A Spearman's rank-order correlation was run to determine the relationship between total time on tasks and familiarity with Google Scholar in the control condition at the $p < .05$ level of significance. There was a weak negative correlation between total time on tasks and familiarity with Google Scholar in the control condition, which was statistically significant ($r_s(8) = -.363, p = .049$).
A Spearman's rank-order correlation was run to determine the relationship between total time on tasks and familiarity with Google Scholar in the experimental condition at the p < .05 level of significance. There was no correlation between total time on tasks and familiarity with Google Scholar in the experimental condition, which was statistically significant ($r_s(8) = -.095, p = .617$).
A Spearman's rank-order correlation was run to determine the relationship between total time on tasks and familiarity with cyberbullying in the control condition at the p < .05 level of significance. There was no correlation between total time on tasks and familiarity with cyberbullying in the control condition, which was statistically significant ($r_s(8) = .036, p = .852$).
A Spearman's rank-order correlation was run to determine the relationship between total time on tasks and familiarity with cyberbullying in the experimental condition at the p < .05 level of significance. There was no correlation between total time on tasks and familiarity with cyberbullying in the experimental condition, which was statistically significant ($r_s(8) = .168, p = .374$).
A Spearman's rank-order correlation was run to determine the relationship between relevancy scores and estimated number of papers written in the control condition at the p < .05 level of significance. There was no correlation between relevancy scores and estimated number of papers written in the control condition, which was statistically significant ($r_s(8) = -.032, p = .865$).
A Spearman's rank-order correlation was run to determine the relationship between relevancy scores and estimated number of papers written in the experimental condition at the $p < .05$ level of significance. There was no correlation between relevancy scores and estimated number of papers written in the experimental condition, which was statistically significant ($r_s(8) = .064$, $p = .736$).
A Spearman's rank-order correlation was run to determine the relationship between relevancy scores and the estimated graduating class size in the control condition at the p < .05 level of significance. There was no correlation between relevancy scores and the estimated graduating class size in the control condition, which was statistically significant ($r_s(8) = -.181$, $p = .338$).
A Spearman's rank-order correlation was run to determine the relationship between relevancy scores and the estimated graduating class size in the experimental condition at the \( p < .05 \) level of significance. There was no correlation between relevancy scores and the estimated graduating class size in the experimental condition, which was statistically significant \( (r_s(8) = -.172, p = .363) \).
A Spearman's rank-order correlation was run to determine the relationship between relevancy scores and the number of credit hours taken in the control condition at the $p < .05$ level of significance. There was no correlation between relevancy scores and the number of credit hours taken in the control condition, which was statistically significant ($r_s(8) = -.118, p = .534$).
A Spearman's rank-order correlation was run to determine the relationship between relevancy scores and the number of credit hours taken in the experimental condition at the $p < .05$ level of significance. There was no correlation between relevancy scores and the number of credit hours taken in the experimental condition, which was statistically significant ($r_s(8) = .085, p = .655$).
A Spearman's rank-order correlation was run to determine the relationship between relevancy scores and familiarity with Google Scholar in the control condition at the p < .05 level of significance. There was no correlation between relevancy scores and familiarity with Google Scholar in the control condition, which was statistically significant ($r_s(8) = .211, p = .264$).
A Spearman's rank-order correlation was run to determine the relationship between relevancy scores and familiarity with Google Scholar in the experimental condition at the $p < .05$ level of significance. There was no correlation between relevancy scores and familiarity with Google Scholar in the experimental condition, which was statistically significant ($r_s(8) = .276, p = .140$).
A Spearman's rank-order correlation was run to determine the relationship between relevancy scores and familiarity with cyberbullying in the control condition at the p < .05 level of significance. There was no correlation between relevancy scores and familiarity with cyberbullying in the control condition, which was statistically significant ($r_s(8) = -.013, p = .944$).
A Spearman's rank-order correlation was run to determine the relationship between relevancy scores and familiarity with cyberbullying in the experimental condition at the $p < .05$ level of significance. There was no correlation between relevancy scores and familiarity with cyberbullying in the experimental condition, which was statistically significant ($r_s(8) = .187, p = .323$).
### APPENDIX F. AOI TITLE TIMES PER PARTICIPANT

<table>
<thead>
<tr>
<th>Participant</th>
<th>AOI Control Mean Time Title</th>
<th>AOI Experimental Mean Time Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participant 1</td>
<td>10.77</td>
<td>19.45</td>
</tr>
<tr>
<td>Participant 2</td>
<td>12.8</td>
<td>16.07</td>
</tr>
<tr>
<td>Participant 3</td>
<td>19.21</td>
<td>12.84</td>
</tr>
<tr>
<td>Participant 4</td>
<td>4.33</td>
<td>2.93</td>
</tr>
<tr>
<td>Participant 5</td>
<td>12.62</td>
<td>16.07</td>
</tr>
<tr>
<td>Participant 6</td>
<td>2.27</td>
<td>1.52</td>
</tr>
<tr>
<td>Participant 7</td>
<td>41.18</td>
<td>35.78</td>
</tr>
<tr>
<td>Participant 8</td>
<td>12.84</td>
<td>2.46</td>
</tr>
<tr>
<td>Participant 9</td>
<td>19.34</td>
<td>10.8</td>
</tr>
<tr>
<td>Participant 10</td>
<td>7.31</td>
<td>1.19</td>
</tr>
<tr>
<td>Participant 11</td>
<td>12.13</td>
<td>2.13</td>
</tr>
<tr>
<td>Participant 12</td>
<td>15.88</td>
<td>18.48</td>
</tr>
<tr>
<td>Participant 13</td>
<td>19.68</td>
<td>9.54</td>
</tr>
<tr>
<td>Participant 14</td>
<td>22.23</td>
<td>7.8</td>
</tr>
<tr>
<td>Participant 15</td>
<td>27.31</td>
<td>23.26</td>
</tr>
<tr>
<td>Participant 16</td>
<td>1.12</td>
<td>8.68</td>
</tr>
<tr>
<td>Participant 17</td>
<td>1.46</td>
<td>2.93</td>
</tr>
<tr>
<td>Participant 18</td>
<td>17.97</td>
<td>27.85</td>
</tr>
<tr>
<td>Participant 19</td>
<td>43.21</td>
<td>45.54</td>
</tr>
<tr>
<td>Participant 20</td>
<td>100.96</td>
<td>37.29</td>
</tr>
<tr>
<td>Participant 21</td>
<td>65.46</td>
<td>17.29</td>
</tr>
<tr>
<td>Participant 22</td>
<td>26.78</td>
<td>19.99</td>
</tr>
<tr>
<td>Participant 23</td>
<td>30.85</td>
<td>12.66</td>
</tr>
<tr>
<td>Participant 24</td>
<td>21.6</td>
<td>8.54</td>
</tr>
<tr>
<td>Participant 25</td>
<td>1.45</td>
<td>7.94</td>
</tr>
<tr>
<td>Participant 26</td>
<td>10.39</td>
<td>7.48</td>
</tr>
<tr>
<td>Participant 27</td>
<td>17.22</td>
<td>16.96</td>
</tr>
<tr>
<td>Participant 28</td>
<td>30.97</td>
<td>43.15</td>
</tr>
<tr>
<td>Participant 29</td>
<td>9.05</td>
<td>14.91</td>
</tr>
<tr>
<td>Participant 30</td>
<td>13.21</td>
<td>6.05</td>
</tr>
<tr>
<td>Participant 31</td>
<td>9.66</td>
<td>13.46</td>
</tr>
<tr>
<td>Participant 32</td>
<td>11.57</td>
<td>20.35</td>
</tr>
</tbody>
</table>
### APPENDIX G. AOI AUTHOR/PUBLISHER TIMES PER PARTICIPANT

<table>
<thead>
<tr>
<th>Participant</th>
<th>AOI Control Mean Time Author/Publisher Seconds</th>
<th>AOI Experimental Mean Time Author/Publisher Seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participant 1</td>
<td>4.32</td>
<td>6.52</td>
</tr>
<tr>
<td>Participant 2</td>
<td>7.35</td>
<td>2.47</td>
</tr>
<tr>
<td>Participant 3</td>
<td>6.18</td>
<td>11.18</td>
</tr>
<tr>
<td>Participant 4</td>
<td>5.43</td>
<td>0.37</td>
</tr>
<tr>
<td>Participant 5</td>
<td>7.61</td>
<td>5.68</td>
</tr>
<tr>
<td>Participant 6</td>
<td>1.15</td>
<td>0</td>
</tr>
<tr>
<td>Participant 7</td>
<td>24.03</td>
<td>16.13</td>
</tr>
<tr>
<td>Participant 8</td>
<td>6.91</td>
<td>1.85</td>
</tr>
<tr>
<td>Participant 9</td>
<td>9.88</td>
<td>2.27</td>
</tr>
<tr>
<td>Participant 10</td>
<td>2.99</td>
<td>1.2</td>
</tr>
<tr>
<td>Participant 11</td>
<td>5.52</td>
<td>3.6</td>
</tr>
<tr>
<td>Participant 12</td>
<td>5.76</td>
<td>6.46</td>
</tr>
<tr>
<td>Participant 13</td>
<td>5.65</td>
<td>10.97</td>
</tr>
<tr>
<td>Participant 14</td>
<td>10.79</td>
<td>1.74</td>
</tr>
<tr>
<td>Participant 15</td>
<td>6.38</td>
<td>6.72</td>
</tr>
<tr>
<td>Participant 16</td>
<td>0.7</td>
<td>0.15</td>
</tr>
<tr>
<td>Participant 17</td>
<td>0.86</td>
<td>4.69</td>
</tr>
<tr>
<td>Participant 18</td>
<td>5.11</td>
<td>8.01</td>
</tr>
<tr>
<td>Participant 19</td>
<td>9.84</td>
<td>22.88</td>
</tr>
<tr>
<td>Participant 20</td>
<td>20.81</td>
<td>31.09</td>
</tr>
<tr>
<td>Participant 21</td>
<td>33.33</td>
<td>9.41</td>
</tr>
<tr>
<td>Participant 22</td>
<td>9.78</td>
<td>21.43</td>
</tr>
<tr>
<td>Participant 23</td>
<td>17.58</td>
<td>6.5</td>
</tr>
<tr>
<td>Participant 24</td>
<td>2.62</td>
<td>9.27</td>
</tr>
<tr>
<td>Participant 25</td>
<td>0.17</td>
<td>1.4</td>
</tr>
<tr>
<td>Participant 26</td>
<td>2.2</td>
<td>0.62</td>
</tr>
<tr>
<td>Participant 27</td>
<td>3.43</td>
<td>3.78</td>
</tr>
<tr>
<td>Participant 28</td>
<td>31.4</td>
<td>33.77</td>
</tr>
<tr>
<td>Participant 29</td>
<td>6.58</td>
<td>5.08</td>
</tr>
<tr>
<td>Participant 30</td>
<td>5.35</td>
<td>1.83</td>
</tr>
<tr>
<td>Participant 31</td>
<td>10.06</td>
<td>3.89</td>
</tr>
<tr>
<td>Participant 32</td>
<td>6.5</td>
<td>8.49</td>
</tr>
</tbody>
</table>
### APPENDIX H. AOI META DESCRIPTION TIMES PER PARTICIPANT

<table>
<thead>
<tr>
<th>Areas of Interest</th>
<th>Mean Gaze Time Meta Description</th>
<th>Participant</th>
<th>AOI Control Mean Time Meta Description Seconds</th>
<th>AOI Experimental Mean Time Meta Description Seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Participant</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>65.98</td>
<td>134.35</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>9.53</td>
<td>15.33</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>9.95</td>
<td>66.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>9.46</td>
<td>5.67</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>6.45</td>
<td>12.88</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>7.7</td>
<td>7.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7</td>
<td>80.64</td>
<td>245.33</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
<td>7.37</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>9</td>
<td>60.43</td>
<td>50.44</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>25.76</td>
<td>24.54</td>
</tr>
<tr>
<td></td>
<td></td>
<td>11</td>
<td>25.61</td>
<td>87.13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12</td>
<td>4.47</td>
<td>33.49</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13</td>
<td>5.34</td>
<td>24.39</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14</td>
<td>66.53</td>
<td>101.67</td>
</tr>
<tr>
<td></td>
<td></td>
<td>15</td>
<td>6.34</td>
<td>8.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>16</td>
<td>0.56</td>
<td>93.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>17</td>
<td>11.34</td>
<td>102.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>18</td>
<td>25.34</td>
<td>199.66</td>
</tr>
<tr>
<td></td>
<td></td>
<td>19</td>
<td>33.64</td>
<td>228.97</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20</td>
<td>68.72</td>
<td>347.49</td>
</tr>
<tr>
<td></td>
<td></td>
<td>21</td>
<td>49.6</td>
<td>273.38</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22</td>
<td>71.82</td>
<td>268.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td>23</td>
<td>58.68</td>
<td>176.61</td>
</tr>
<tr>
<td></td>
<td></td>
<td>24</td>
<td>7.64</td>
<td>126.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25</td>
<td>1.31</td>
<td>35.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>26</td>
<td>3.89</td>
<td>6.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td>27</td>
<td>3.57</td>
<td>10.69</td>
</tr>
<tr>
<td></td>
<td></td>
<td>28</td>
<td>59.9</td>
<td>249.37</td>
</tr>
<tr>
<td></td>
<td></td>
<td>29</td>
<td>20.4</td>
<td>188.99</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
<td>20.98</td>
<td>9.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>31</td>
<td>41.03</td>
<td>101.46</td>
</tr>
<tr>
<td></td>
<td></td>
<td>32</td>
<td>25.3</td>
<td>145.24</td>
</tr>
</tbody>
</table>
APPENDIX I. AOI CITED TIMES PER PARTICIPANT

<table>
<thead>
<tr>
<th>Participant</th>
<th>AOI Control Mean Time Cited</th>
<th>AOI Experimental Mean Time Cited</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participant 1</td>
<td>13.7</td>
<td>6.23</td>
</tr>
<tr>
<td>Participant 2</td>
<td>0.75</td>
<td>0.88</td>
</tr>
<tr>
<td>Participant 3</td>
<td>6.18</td>
<td>5.77</td>
</tr>
<tr>
<td>Participant 4</td>
<td>5.43</td>
<td>1.55</td>
</tr>
<tr>
<td>Participant 5</td>
<td>7.61</td>
<td>1.71</td>
</tr>
<tr>
<td>Participant 6</td>
<td>1.15</td>
<td>5.6</td>
</tr>
<tr>
<td>Participant 7</td>
<td>24.03</td>
<td>3.87</td>
</tr>
<tr>
<td>Participant 8</td>
<td>6.91</td>
<td>1.18</td>
</tr>
<tr>
<td>Participant 9</td>
<td>8.45</td>
<td>2.03</td>
</tr>
<tr>
<td>Participant 10</td>
<td>0.92</td>
<td>0</td>
</tr>
<tr>
<td>Participant 11</td>
<td>0.5</td>
<td>0.76</td>
</tr>
<tr>
<td>Participant 12</td>
<td>0.75</td>
<td>1.4</td>
</tr>
<tr>
<td>Participant 13</td>
<td>2.83</td>
<td>0.5</td>
</tr>
<tr>
<td>Participant 14</td>
<td>6.48</td>
<td>0.95</td>
</tr>
<tr>
<td>Participant 15</td>
<td>1.75</td>
<td>2.68</td>
</tr>
<tr>
<td>Participant 16</td>
<td>0.09</td>
<td>5.24</td>
</tr>
<tr>
<td>Participant 17</td>
<td>2.59</td>
<td>3.43</td>
</tr>
<tr>
<td>Participant 18</td>
<td>4.08</td>
<td>4.02</td>
</tr>
<tr>
<td>Participant 19</td>
<td>3.7</td>
<td>13.43</td>
</tr>
<tr>
<td>Participant 20</td>
<td>6.48</td>
<td>10.69</td>
</tr>
<tr>
<td>Participant 21</td>
<td>12.02</td>
<td>9.1</td>
</tr>
<tr>
<td>Participant 22</td>
<td>4.02</td>
<td>3.08</td>
</tr>
<tr>
<td>Participant 23</td>
<td>1.59</td>
<td>4.48</td>
</tr>
<tr>
<td>Participant 24</td>
<td>0.62</td>
<td>3.82</td>
</tr>
<tr>
<td>Participant 25</td>
<td>3.54</td>
<td>9.81</td>
</tr>
<tr>
<td>Participant 26</td>
<td>3.95</td>
<td>3.71</td>
</tr>
<tr>
<td>Participant 27</td>
<td>2.32</td>
<td>0.65</td>
</tr>
<tr>
<td>Participant 28</td>
<td>7.98</td>
<td>6.1</td>
</tr>
<tr>
<td>Participant 29</td>
<td>12.73</td>
<td>1.33</td>
</tr>
<tr>
<td>Participant 30</td>
<td>4.3</td>
<td>2.09</td>
</tr>
<tr>
<td>Participant 31</td>
<td>4</td>
<td>6.08</td>
</tr>
<tr>
<td>Participant 32</td>
<td>4.7</td>
<td>2.07</td>
</tr>
</tbody>
</table>
APPENDIX J. NUMBER OF SURROGATE DOCUMENT TRANSITIONS

<table>
<thead>
<tr>
<th>Participant</th>
<th>AOI Control Document Transitions</th>
<th>AOI Title Document Transitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participant 1</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Participant 2</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>Participant 3</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Participant 4</td>
<td>13</td>
<td>10</td>
</tr>
<tr>
<td>Participant 5</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Participant 6</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Participant 7</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Participant 8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Participant 9</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>Participant 10</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>Participant 11</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>Participant 12</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Participant 13</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Participant 14</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Participant 15</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Participant 16</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>Participant 17</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Participant 18</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Participant 19</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Participant 20</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Participant 21</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Participant 22</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Participant 23</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Participant 24</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Participant 25</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Participant 26</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>Participant 27</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>Participant 28</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Participant 29</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Participant 30</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>Participant 31</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>Participant 32</td>
<td>9</td>
<td>0</td>
</tr>
</tbody>
</table>
### APPENDIX K. TOTAL TIME ON TASK PER PARTICIPANT

<table>
<thead>
<tr>
<th>Total Time on Task</th>
<th>Control Time on Task</th>
<th>Experimental Time on Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participant 1</td>
<td>04:27.5</td>
<td>05:57.1</td>
</tr>
<tr>
<td>Participant 2</td>
<td>06:36.9</td>
<td>05:52.3</td>
</tr>
<tr>
<td>Participant 3</td>
<td>01:39.9</td>
<td>08:11.5</td>
</tr>
<tr>
<td>Participant 4</td>
<td>23:01.5</td>
<td>12:33.2</td>
</tr>
<tr>
<td>Participant 5</td>
<td>07:32.5</td>
<td>07:45.1</td>
</tr>
<tr>
<td>Participant 6</td>
<td>04:54.0</td>
<td>18:25.3</td>
</tr>
<tr>
<td>Participant 7</td>
<td>05:46.0</td>
<td>08:02.7</td>
</tr>
<tr>
<td>Participant 8</td>
<td>19:21.2</td>
<td>06:50.1</td>
</tr>
<tr>
<td>Participant 9</td>
<td>08:07.1</td>
<td>02:21.8</td>
</tr>
<tr>
<td>Participant 10</td>
<td>08:54.8</td>
<td>05:01.6</td>
</tr>
<tr>
<td>Participant 11</td>
<td>04:42.5</td>
<td>02:40.9</td>
</tr>
<tr>
<td>Participant 12</td>
<td>05:29.2</td>
<td>04:04.7</td>
</tr>
<tr>
<td>Participant 13</td>
<td>16:02.8</td>
<td>13:25.3</td>
</tr>
<tr>
<td>Participant 14</td>
<td>05:36.0</td>
<td>03:04.1</td>
</tr>
<tr>
<td>Participant 15</td>
<td>11:05.7</td>
<td>08:17.4</td>
</tr>
<tr>
<td>Participant 16</td>
<td>05:16.5</td>
<td>05:13.4</td>
</tr>
<tr>
<td>Participant 17</td>
<td>06:22.3</td>
<td>06:08.6</td>
</tr>
<tr>
<td>Participant 18</td>
<td>04:57.3</td>
<td>06:16.6</td>
</tr>
<tr>
<td>Participant 19</td>
<td>03:36.6</td>
<td>08:44.2</td>
</tr>
<tr>
<td>Participant 20</td>
<td>13:33.1</td>
<td>09:58.1</td>
</tr>
<tr>
<td>Participant 21</td>
<td>07:40.3</td>
<td>05:45.1</td>
</tr>
<tr>
<td>Participant 22</td>
<td>07:22.0</td>
<td>04:33.1</td>
</tr>
<tr>
<td>Participant 23</td>
<td>10:05.5</td>
<td>06:25.3</td>
</tr>
<tr>
<td>Participant 24</td>
<td>04:51.1</td>
<td>04:19.0</td>
</tr>
<tr>
<td>Participant 25</td>
<td>04:07.4</td>
<td>05:20.4</td>
</tr>
<tr>
<td>Participant 26</td>
<td>06:38.9</td>
<td>06:56.6</td>
</tr>
<tr>
<td>Participant 27</td>
<td>10:32.0</td>
<td>11:21.5</td>
</tr>
<tr>
<td>Participant 28</td>
<td>06:53.6</td>
<td>08:44.0</td>
</tr>
<tr>
<td>Participant 29</td>
<td>01:30.5</td>
<td>04:41.5</td>
</tr>
<tr>
<td>Participant 30</td>
<td>04:44.7</td>
<td>06:40.5</td>
</tr>
<tr>
<td>Participant 31</td>
<td>08:04.0</td>
<td>08:31.2</td>
</tr>
<tr>
<td>Participant 32</td>
<td>06:12.9</td>
<td>05:23.1</td>
</tr>
</tbody>
</table>
### APPENDIX L. RELEVANCY ACCURACY PER PARTICIPANT

<table>
<thead>
<tr>
<th>Participant</th>
<th>Control Relevancy Accuracy</th>
<th>Experimental Relevancy Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participant 1</td>
<td>70</td>
<td>60</td>
</tr>
<tr>
<td>Participant 2</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>Participant 3</td>
<td>100</td>
<td>90</td>
</tr>
<tr>
<td>Participant 4</td>
<td>90</td>
<td>80</td>
</tr>
<tr>
<td>Participant 5</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>Participant 6</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>Participant 7</td>
<td>100</td>
<td>90</td>
</tr>
<tr>
<td>Participant 8</td>
<td>10</td>
<td>40</td>
</tr>
<tr>
<td>Participant 9</td>
<td>70</td>
<td>90</td>
</tr>
<tr>
<td>Participant 10</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>Participant 11</td>
<td>70</td>
<td>80</td>
</tr>
<tr>
<td>Participant 12</td>
<td>80</td>
<td>90</td>
</tr>
<tr>
<td>Participant 13</td>
<td>90</td>
<td>70</td>
</tr>
<tr>
<td>Participant 14</td>
<td>90</td>
<td>80</td>
</tr>
<tr>
<td>Participant 15</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>Participant 16</td>
<td>80</td>
<td>90</td>
</tr>
<tr>
<td>Participant 17</td>
<td>90</td>
<td>100</td>
</tr>
<tr>
<td>Participant 18</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>Participant 19</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>Participant 20</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>Participant 21</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>Participant 22</td>
<td>80</td>
<td>90</td>
</tr>
<tr>
<td>Participant 23</td>
<td>70</td>
<td>70</td>
</tr>
<tr>
<td>Participant 24</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>Participant 25</td>
<td>90</td>
<td>70</td>
</tr>
<tr>
<td>Participant 26</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>Participant 27</td>
<td>100</td>
<td>90</td>
</tr>
<tr>
<td>Participant 28</td>
<td>90</td>
<td>80</td>
</tr>
<tr>
<td>Participant 29</td>
<td>90</td>
<td>80</td>
</tr>
<tr>
<td>Participant 30</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>Participant 31</td>
<td>90</td>
<td>80</td>
</tr>
<tr>
<td>Participant 32</td>
<td>90</td>
<td>100</td>
</tr>
</tbody>
</table>
APPENDIX M. SURVEY RESPONSES

Question 1
*How old are you?*

<table>
<thead>
<tr>
<th>Answer</th>
<th>Response</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>17-18</td>
<td>2</td>
<td>7%</td>
</tr>
<tr>
<td>18-19</td>
<td>19</td>
<td>63%</td>
</tr>
<tr>
<td>20-21</td>
<td>4</td>
<td>13%</td>
</tr>
<tr>
<td>21-older</td>
<td>5</td>
<td>17%</td>
</tr>
</tbody>
</table>

Question 2
*What is your gender?*

<table>
<thead>
<tr>
<th>Answer</th>
<th>Response</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>17</td>
<td>57%</td>
</tr>
<tr>
<td>Female</td>
<td>13</td>
<td>43%</td>
</tr>
<tr>
<td>Transgender</td>
<td>0</td>
<td>0%</td>
</tr>
</tbody>
</table>

Question 3
*Approximately (best estimate) how papers have you written for your college classes?*

<table>
<thead>
<tr>
<th>Answer</th>
<th>Response</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-6</td>
<td>7</td>
<td>23%</td>
</tr>
<tr>
<td>7-9</td>
<td>11</td>
<td>37%</td>
</tr>
<tr>
<td>10-12</td>
<td>4</td>
<td>13%</td>
</tr>
<tr>
<td>13-15</td>
<td>3</td>
<td>10%</td>
</tr>
<tr>
<td>16 or more</td>
<td>5</td>
<td>17%</td>
</tr>
</tbody>
</table>

Question 4
*How many people were in your graduating class?*

<table>
<thead>
<tr>
<th>Answer</th>
<th>Response</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than 100</td>
<td>4</td>
<td>12%</td>
</tr>
<tr>
<td>101-300</td>
<td>6</td>
<td>21%</td>
</tr>
<tr>
<td>301-500</td>
<td>10</td>
<td>34%</td>
</tr>
<tr>
<td>501-1000</td>
<td>8</td>
<td>28%</td>
</tr>
<tr>
<td>1000 or more</td>
<td>1</td>
<td>3%</td>
</tr>
</tbody>
</table>
Question 5

*Approximately (best estimate) how many credit hours have you completed, (not including this semester)?*

<table>
<thead>
<tr>
<th>Answer</th>
<th>Response</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than 12</td>
<td>1</td>
<td>3%</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>3%</td>
</tr>
<tr>
<td>15</td>
<td>10</td>
<td>33%</td>
</tr>
<tr>
<td>18</td>
<td>5</td>
<td>17%</td>
</tr>
<tr>
<td>21 or more</td>
<td>13</td>
<td>43%</td>
</tr>
</tbody>
</table>

Question 6

*How familiar are you with Google Scholar?*

<table>
<thead>
<tr>
<th>Answer</th>
<th>Response</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>I have never used Google Scholar before.</td>
<td>11</td>
<td>37%</td>
</tr>
<tr>
<td>I have used Google Scholar once or twice.</td>
<td>11</td>
<td>37%</td>
</tr>
<tr>
<td>I have used Google Scholar several times before.</td>
<td>7</td>
<td>23%</td>
</tr>
<tr>
<td>I use Google Scholar often.</td>
<td>1</td>
<td>3%</td>
</tr>
<tr>
<td>I almost always use Google Scholar.</td>
<td>0</td>
<td>0%</td>
</tr>
</tbody>
</table>

Question 7

*Did you already know allot about the topic of cyberbullying?*

<table>
<thead>
<tr>
<th>Answer</th>
<th>Response</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>I have heard of it, but I have never paid much attention to it.</td>
<td>2</td>
<td>7%</td>
</tr>
<tr>
<td>I know a little about it, but I don’t know more than the basics.</td>
<td>4</td>
<td>13%</td>
</tr>
<tr>
<td>I know some about it from the news and classes, enough that I could explain it to some one else.</td>
<td>18</td>
<td>60%</td>
</tr>
<tr>
<td>I know more than most people from the news, classes, and my own interest in the topic.</td>
<td>5</td>
<td>17%</td>
</tr>
<tr>
<td>I know allot about the topic, I have read articles, watched videos, or been to lectures about it.</td>
<td>1</td>
<td>3%</td>
</tr>
</tbody>
</table>
Question 8  
**Did you feel like you were faster in one version of Google Scholar or another?**

<table>
<thead>
<tr>
<th>Answer</th>
<th>Response</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neither was faster.</td>
<td>6</td>
<td>20%</td>
</tr>
<tr>
<td>I felt faster in the version with shorter descriptions.</td>
<td>9</td>
<td>30%</td>
</tr>
<tr>
<td>I felt faster in the version with longer descriptions.</td>
<td>15</td>
<td>50%</td>
</tr>
</tbody>
</table>

Question 9  
**Of the two versions of Google Scholar did you like one better than the other?**

<table>
<thead>
<tr>
<th>Answer</th>
<th>Response</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>I have no preference.</td>
<td>6</td>
<td>20%</td>
</tr>
<tr>
<td>I prefer the one with shorter descriptions.</td>
<td>5</td>
<td>17%</td>
</tr>
<tr>
<td>I prefer the one with longer descriptions.</td>
<td>19</td>
<td>63%</td>
</tr>
</tbody>
</table>

Question 10  
**If you had to use Google Scholar for a class research paper would you prefer one version over the other?**

<table>
<thead>
<tr>
<th>Answer</th>
<th>Response</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>I have no preference.</td>
<td>2</td>
<td>7%</td>
</tr>
<tr>
<td>I would rather use the one with shorter descriptions.</td>
<td>7</td>
<td>23%</td>
</tr>
<tr>
<td>I would rather use the one with longer descriptions.</td>
<td>21</td>
<td>70%</td>
</tr>
</tbody>
</table>

Question 11  
**Did you remember feeling more confident with one version of Google Scholar than the other?**

<table>
<thead>
<tr>
<th>Answer</th>
<th>Response</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>I don't remember feeling more confident with either version.</td>
<td>11</td>
<td>37%</td>
</tr>
<tr>
<td>I felt more confident with the shorter descriptions.</td>
<td>4</td>
<td>13%</td>
</tr>
<tr>
<td>I felt more confident with the longer descriptions.</td>
<td>15</td>
<td>50%</td>
</tr>
</tbody>
</table>
Question 12

*Do you remember opening other webpages or PDF documents more often in one environment than the other?*

<table>
<thead>
<tr>
<th>Answer</th>
<th>Response</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>I don't remember a difference with either version.</td>
<td>8</td>
<td>27%</td>
</tr>
<tr>
<td>I opened more with the shorter descriptions.</td>
<td>17</td>
<td>57%</td>
</tr>
<tr>
<td>I opened more with the longer descriptions.</td>
<td>5</td>
<td>17%</td>
</tr>
</tbody>
</table>
VITA

In 1978 Nathan Lowrance was born as the first of two children to John and Heidi Lowrance. He was raised in a progressive household, exposed to travel, language, and the arts at an early age. His father’s interest in science fiction had the largest effect on his academic life, exposing him to a form of literature that asked the big questions about human nature and humanity’s role in the world. This, combined with his mother’s dedication to travel and connecting to her Swiss heritage, helped shape Nathan’s connection to the world outside the US and formed his interest in where humanity as a whole was heading.

After high school, Nathan had the good fortune to join the Rotary exchange program and lived for a year in Germany, where he attended Gymnasium. This experience allowed him to make many lasting friendships, learn, and grow. While there Nathan spent much of his time in museums and art galleries, and at every opportunity explored many of Europe’s great cities. He returned to the United States to pursue undergraduate degrees in philosophy and German at Drury University, graduating with honors in 2002. After spending time working a variety of jobs, including teaching high school and working in a drug and alcohol treatment center, Nathan returned to higher education seeking a master’s degree in Information Sciences and Learning Technologies from the University of Missouri, graduating in May 2011. After this, Nathan continued his academic pursuits in studying human computer interaction, human information behavior and cognitive biases. Throughout his time at the University of Missouri, Nathan had the opportunity to work on improving usability for a variety of electronic health records applications.

Nathan continues to explore human information behavior and medical health care applications. When not relaxing in the company of food, family, and friends, Nathan can often be found, exploring meditation practices, watching film, or reading a book.