2007

HighBrow: A context enabled highlighting browser

Ronald J. Zucker
University of South Florida

HighBrow: A Context Enabled Highlighting Browser

by

Ronald J. Zucker

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Computer Science and Engineering
College of Engineering
University of South Florida

Major Professor: Dewey Rundus, Ph.D.
Grisselle Centeno, Ph.D.
Dmitry Goldgof, Ph.D.
Constance Hines, Ph.D.
Rafael Perez, Ph.D.

Date of Approval:
March 26, 2007

Keywords: Digital Annotation, Reading, Cognition, Human Computer Interaction, User Interface

Note to Reader

Note to Reader: the original of this document contains color that is necessary for understanding the data. The original dissertation is on file with the USF library in Tampa, Florida

Dedication

This dissertation is dedicated to my best friend and wife, who sacrificed so much to make this possible. Jeanne allowed us to put our lives on hold while I pursued a lifelong personal dream.

Acknowledgements

I would like to express my gratitude to my advisor Dr. Dewey Rundus for his help and incredible patience. He provided me with valuable insights, encouragement, advice, and humor throughout my Ph.D. studies. I would like to thank my committee: Dr. Grisselle Centeno, Dr. Dmitry Goldgof, Dr. Constance Hines, and Dr. Rafael Perez. I would also like to thank my fellow students at the University of South Florida who worked as a team to encourage, support, and challenge me. These include but are not limited to Soumyaroop Roy, Ana Staninska, Isabela Moura, Matt Long, and Pedrow Whitehouse. My colleagues at the University of North Florida, Drs. Behrooz Abassi, Dominik Guess, Jeff Michaelman, Donna Mohr, Dan Philip, Bob Roggio, and Yap Chua all provided support or volunteered to have their students participate in the experiments. Pat Nelson reviewed the dissertation format. Greg Blajian reviewed and tested the software to ensure bug free experiments. Albert Ritzhaupt provided support and supplied students for the pilot experiment. Last, but not least, I would like to acknowledge all of the 150 plus participants who took part in this study.

Table of Contents

List of Tables

List of Figures

HighBrow: A Context Enabled Highlighting Browser

Ronald J. Zucker

ABSTRACT

As the World Wide Web continues to grow, more and more information is retrieved online. A person visiting a Web site has a choice of whether to skip, skim, deep read, bookmark for later revisiting, print a document, or any combination of these options. Recently, several tools have been developed to allow users to digitally annotate Web documents. These tools allow users to highlight, make text notes, and scrape information (a form of copy and paste). This dissertation introduces a new form of annotating, called context highlighting. Context highlighting is the ability to mark the text that surrounds keywords or phrases. To test the benefits and costs involved in keyword and context highlighting, a prototype browser called HighBrow was developed specifically for this dissertation that is capable of highlighting both keywords and the supporting context.

The first experiment in this dissertation addresses possible benefits of highlighting both keywords and context, with respect to improved cognition, ease of use, and likeability for the active reader. The results of this experiment showed promise in cognition, however statistical significance was not achieved. Participants liked HighBrow, finding it easy to learn and use. The context/keywords highlighters produced

significantly smaller keyword phrases than the keyword only highlighters and the amount of time required to do the additional highlighting was not considered detrimental.

When active readers highlight keywords and phrases as well as the surrounding context, HighBrow will produce a context summary. A second experiment was conducted to show that passive readers of a context/keyword summary will be more efficient by reducing preparation time and scoring as well as readers who read the entire document or a keyword only summary. A third experiment was conducted to determine if patterns of highlighting changed over time. The results of the third experiment were disappointing as too many students opted out, providing too little data to make any conclusions.

Overall, context highlighting has potential with respect to cognition for both active and passive readers and reducing preparation time for passive readers.

Chapter One:

Introduction

This chapter introduces the concept of context highlighting and provides the motivation for, and questions regarding, this new form of annotation. The chapter concludes with a brief description of how the remaining chapters of the dissertation are organized.

1.1 Context Highlighting

Highlighting is a method of annotating a document that is usually used to signal future attention, to help mark important places, aid memory, and trace progress through difficult narrative (Marshall, 1997). Simple highlighting involves a marker or mouse that uses real or digital ink. The reader simply moves a marker or mouse over the appropriate material and the material is highlighted.

For digital documents, the act of highlighting can be followed by a copy and paste operation to create a summary of the digital document. The paper and ink document may be similarly organized by copying the original document, then physically cutting the highlighted context and pasting it elsewhere (Marshall, 1998), a very tedious and time-consuming process. Digital highlights can be used to quickly create indices of the keywords or phrases (Brown & Brown, 2004). Digital highlights can be retrieved easily and since these highlights have a reference to the location in the full document (Turney, 1999), the reader can locate the original context surrounding the keyword or phrase with much less effort than the reader of a paper document. Other advantages of digital highlighting are the ability to hide or show the highlights on demand and to modify or delete existing highlights (Cousins, Baldonado, & Paepcke, 2000). Digital highlights can easily be organized in a number of ways, for example: alphabetically or by category (Brown & Brown, 2004).

Context highlighting is a new concept that has evolved from simple highlighting, allowing the reader to highlight not only keywords or phrases but also to highlight the surrounding context. The main purpose of context highlighting would be to assist in interpretation or understanding of the keyword or phrase. Context highlighting can be done in a non-digital way by selecting a keyword color marker (e.g., bright yellow) and a context color marker (e.g., light pink).

With digital documents, when the reader selects the surrounding context, the software can create a summary, called a context summary, which will be much more readable and understandable and still maintain the highlighted keywords or phrases. The original digital document will also show the context and keywords or phrases allowing the reader to modify the context as the need arises, adding more context if needed, or reducing the context if too much material has been highlighted.

Allowing one to highlight context as well as keywords helps promote deeper cognitive processing since importance is being filtered at two levels: the key concept and the supporting information. Context highlighting forces the reader to reread the passage in order to provide adequate context for the keyword or phrase.

With respect to document sizes (in words), the size of a context summary lies somewhere between the size of the full document and the size of a keyword summary which consists solely of keywords or phrases. The reduced size of the context summary versus the full document helps to reduce the time required to read the material. While the keyword summary is smaller than the context summary, the amount of time spent reading, and perhaps trying to comprehend, the keyword summary is likely to be only marginally shorter.

Readers who are not highlighters might also benefit from the context summaries created by other annotating readers. These context summaries would allow readers to concentrate only on the information that the original annotators thought was important and require less time to understand the content. Since the context summary contains supporting text, it is more readable than the keyword summary and it is more informative. Emphasizing importance within the summary through the use of keyword highlighting also helps draw attention to the key concepts within the document.

If a reader can recall an equal amount of important information in less time, the reader may be considered more efficient. Thus if a reader of a context summary can spend less time reading the context summary than reading a full document, and score as well as or better than reading the entire document, one can say that the process is more efficient.

1.2 Motivation

Reading documents can be a time consuming process. If the document contains a lot of information, the use of annotations may be necessary to help recall and summarize the content (Denoue & Vignollet, 2000). When people annotate using highlighting, their behaviors vary. In some cases, the reader may use multiple colors to categorize the material (Marshall, 1997). The reader may highlight only keywords or phrases, as a way of signaling for future attention, placemarking, and aiding memory, or may elect to highlight large passages of the document in order to mark progress through difficult material (Marshall, 1997). The amount of text that should be highlighted remains an open question (Ostler, 1999), however Crystal, Kubala, and MacIntyre (1999) concluded that annotation quantity is the most important factor for increasing model performance. However, the more text is highlighted and extracted, the more difficult it becomes to recall what was important. For example, Figure 1.1 contains an actual passage, highlighted and extracted from *Expected, Sensed, and Desired: A Framework for Designing Sensing-Based Interaction* (Benford et al., 2005) to create a summary. Reading from this extract and trying to understand what was important about this passage may be time consuming and/or impossible.

> "There is a long and extensive history of user-centered design methods in HCI, including task-analysis techniques that draw on cognitive psychology in order to understand how individuals plan and carry out detailed interactions with particular interfaces, for example, GOMs (John 1996), the use of ethnography to inform system design with an understanding of the social and situated use of technologies in particular environments (Hughes 1992) and participatory design methods that directly involve users as partners in the design process, sometimes through working with low-tech physical prototypes (e.g., Ehn (1991))."
>
> Taken from: *Expected, Sensed, and Desired: A Framework for Designing Sensing-Based Interaction* (Benford et al., 2005)

Figure 1.1 - Extracted Passage without Highlighted Keyword

Context highlighting solves both problems. Context highlighting allows readers to highlight keywords or phrases, keeping the highlights brief and high in importance, and highlight context in order to provide a summary with sufficient content to make it readable. Readers reading the full document will see context highlighting in a lighter shade of yellow and keywords in bright yellow. Yellow was chosen because readers of paper documents often use yellow markers to make passages visually prominent, hence easier to find (Phelps & Wilensky, 1997). Even in the full document, the context highlighting will draw attention to the context as well as the keyword or phrase. The resulting context summary also contains the highlighted keywords or phrases; however, the context is no longer highlighted since the summary, by definition, consists of context (with embedded keywords).

Figure 1.2 is a key phrase highlighted passage, taken from *Annotating the Web: An Exploratory Study of Web Users' Needs for Personal Annotation Tools* (Fu, Ciszek, Marchionini, & Solomon, 2005). One can easily identify the part of the passage that the annotating reader thought was important. The reader wished to emphasize the "incredible convenience of link making" but this phrase taken out of context has little or no meaning. Adding the surrounding context allows the reader to place this phrase in the proper environment with supporting examples.

Any distraction from the reading can be considered a cognitive interrupt. Examples of cognitive interrupts for an annotating reader include: picking up a marker, grabbing the mouse, delays in the process (running out of ink, finding markers, slow response time during the highlighting action), changing from marker to pen or pencil, or changing from mouse to keyboard. Some existing software allows the reader to add annotations, usually as text using a digital form of the Post-It®, in order to explain something about the highlight (Shilman & Wei, 2004; Roscheisen, Mogensen, & Winogred, 1997). This shift between reading, highlighting, and typing text can cause a cognitive interrupt since the reader has changed roles from reader to writer (Marshall, 1998). Research has shown that most persons who highlight do not take time to make textual annotations (Denoue, 2000; Marshall & Brush, 2002).

> "The invention of hypertext technology and supplementary Web technologies provides incredible convenience for link making and path building between existing documents or document snippets. To harness the power of new technologies and give electronic documents some of the same note-taking possibilities as paper documents, people have developed various kinds of annotation tools and applications, from lightweight functions of adding a Web page to a 'Favorites' list or creating a short cut to the page on the toolbar, to a variety of more complex and specific-purpose annotation systems."
>
> Taken from: *Annotating the Web: An Exploratory Study of Web Users' Needs for Personal Annotation Tools* (Fu et al., 2005)

Figure 1.2 - Extracted Passage with Key Phrase Highlighted

1.3 Research Question

This dissertation addresses the following research question:

Does context highlighting improve test preparation and performance?

Since there may be two groups of readers: active readers (also called annotators), readers who read and highlight; passive readers, readers who read previously highlighted materials, this dissertation also includes the following questions:

- Active readers:
    - *Does context and keyword highlighting influence test scores compared to keyword only highlighting?*
    - *Are the number of words of the highlighted portion affected by context highlighting?*

    - *Is the context/keyword highlighter usable and do people like it?*
    - *With use, will the amount of highlighting change? If the users highlighted only keywords in the past will they begin to highlight more words from the text to preserve context?*
    - *Will users voluntarily use the context browser?*
- Passive readers:
    - *Does reading a complete document, a document with context/keywords, or a document with keywords only improve test scores?*
    - *Does reading a complete document, a document with context/keywords, or a document with keywords only reduce study time?*
    - *Is study time and test performance together enhanced by reading a complete document, a document with context/keywords, or a document with keywords only?*

The reason it is important to look at the active readers is that they are the producers of summaries. If there is no advantage or benefit (real or perceived) to the active reader, they will not produce a summary for others to use. People will do things that benefit them even though they do not like them (e.g., flu shots); they also do things that they like doing even if it is not beneficial to them (e.g., smoke cigarettes). Clearly, people will highlight context if they like it and they find it beneficial.

It is also important to look at highlighting patterns over time. Because context highlighting is new, people may alter their highlighting patterns: highlighting more text for the context and less for the keywords or possibly changing the keywords themselves.

Passive readers may benefit from the highlighting of others. If the summary effectively reduces the size while retaining the key content of a document, then the reader will be able to understand the key points of a document without the distraction of unimportant information and without the need to read the entire document. Only if the reader will need to concentrate on the details will there be any need to read the original document. Reviewing the summary may, however, actually encourage a person to read the entire document even if they had never intended to do so.

1.4 Contributions

This dissertation introduces a new tool for reading and remembering important information from a Web document. This approach is new in that context highlighting has never been applied to documents (paper and ink or digitally) in order to build a summary. The main contributions of this work are:

- Introduces a single, simple scheme for highlighting keywords to aid recall and highlighting context to aid in interpretation.
- Demonstrates that active readers, also called annotators, find context highlighting easy to use and enjoy context highlighting and produce useful summaries.
- Provides evidence that passive readers may benefit from the context summaries with respect to efficiency, defined as test score over time to study.
- Development of a prototype browser called HighBrow (from highlighting browser) which enables readers to highlight keywords and context. Literature indicates that the key factors for a computer based annotation scheme to be

successful are speed and simplicity. HighBrow addresses: speed through caching and overlays; simplicity through a simple scheme requiring a mouse only interface which does not require typing. The prototype HighBrow is a very useful tool for developing future studies in highlighting behavior and performance related to highlighting.

1.5 Organization of the Dissertation

The chapters in the dissertation are organized as follows. Chapter Two provides a look at prior work with respect to the broad field of annotation, including why we annotate. The dimensions and systems view of annotations are discussed and the cognitive effects of annotation are addressed. Some of the better-known digital annotation/highlighting systems are also described.

Chapter Three introduces HighBrow, the context enabled highlighting browser, describing its architecture and how it relates to existing architectures. The user interface and the context summary is described and justified. The dimensions and systems view introduced in Chapter Two are related to HighBrow. This chapter also includes a walkthrough of the highlighting process and a brief description of the software testing scenarios for HighBrow. Chapter Three concludes with the various versions of HighBrow used in the experiments presented in Chapter Four.

Chapter Four describes the three experiments performed to answer the questions, including the participants, procedures, tools, and results. Chapter Five is a discussion and summary relating to the experiments described in Chapter Four. Chapter Six contains an overall summary of the dissertation and future work.

Chapter Two:

Related Work

This chapter provides a review of current literature beginning with the broad concept of annotation then narrowing the focus to annotation using highlighting. In addition, this chapter provides an overview of existing work with Web based annotation and more specifically highlighting text on the Web.

2.1 Definition of Annotation

Merriam Webster Online (Definition of annotation – merriam-webster online dictionary.) defines annotation as “a note added by way of comment or explanation”. Marshall (1998) extended the definition to include marginalia, writing between the lines, highlighting, underlining, circling, boxing, and symbolic notation. Cousins, Baldonado, and Paepcke (2000) added “assigning metadata to a literary work”, also pointing out that the move towards digital documents makes defining annotation difficult. Cousins, Baldonado, and Paepcke also revised the definition of annotation to say, “an Annotation is a commentary of an object that: the annotator intends to be separable from the object itself; the reader interprets to be separable from the object itself”. Objects have replaced documents since, in the digital world, annotation is no longer limited to text documents, but can be expanded to multimedia objects (Bottoni et al., 2005)

2.2 Why We Annotate

Annotations are used “to remember, to think, to clarify, and to share” (Ovsiannikov, Arbib & McNeill 1999). According to Marshall (1997), we annotate for many reasons: as a signal for future attention, placemarking, aiding memory, working out problems, interpretation, as a trace of reading through difficult narrative, and incidental reflection. Marshall also points out that annotation takes different forms for different purposes. For example: highlighting or underlining structure, as in topic headings, is used to signal future attention; short highlighting is used for placemarking and aiding memory; extended highlighting and underlining is used to trace reading through difficult narrative. Marginalia and notes are used for problem working and interpretation, though some annotations such as notes, doodles, and drawings that are unrelated to the materials themselves indicate incidental reflection during the reading. Table 2.1 shows the different classes of annotation, the various forms each class may take and the function or usage of the form.

The dominant forms of annotation for Web documents are: text selection and emphasis (including highlighting, underlining, circling, drawing symbols (e.g., stars, asterisks, etc.), association building (e.g., writing notes, drawing sketches), and document re-segmentation (e.g., restructuring or reorganizing the document to fit the reader’s needs)(Fu et al., 2005).

Table 2.1 - Classes of Annotation (Marshall, 1977)

Class	Form	Function
Highlighting	Higher level structures	Procedure signaling for future attention
	Short highlighting	Placemarking and aiding memory
	Extended highlighting or underlining	Tracing progress through difficult narrative
Marking	Within text markings	Placemarking and aiding memory
Marginal annotation	Telegraphic marginal symbols	Procedure signaling for future attention
	Marginal markings	Placemarking and aiding memory
	Marginal notation	Problem-working
Textual annotation	Notation near figures or equations	Problem-working
	Short notes in the margins	Interpretation
	Longer notes and other textual interstices	Interpretation
	Words or phrases between lines of text	Interpretation
	Notes: drawings and other such markings unrelated to the materials themselves	Incidental reflection of the material circumstances of reading

2.3 Cognitive Effects of Annotation

The cognitive effects of pen and ink annotation have been studied for many years. There have been many studies with differing hypothesis on the effects and the reasons for highlighting. Hershberger (1964) studied the effects of highlighting and determined that highlighting did not increase learning enrichment in contrast to the von Restorff, or isolation effect, that predicted that if something "stands out like a sore thumb", for example by the use of color, then the reader will be more likely to remember the item. Wade and Trathen (1989) suggested that it was the ability to distinguish importance not the noting (defined as "any overt study method subjects engage in such as underlining, highlighting, and note taking.") that assisted learning and that noting may be "an epiphenomenon". They did point out in their discussion that noting was not useful in their study, but were also quick to suggest that further study should be done. Peterson's study (1992) showed that students were more apt to answer a question correctly when they had highlighted the relevant words or phrases than if they had not. Nist and Hogrebe (1987) reasoned that underlining text enabled the student to process information at deeper levels.

Annotations should be constructed carefully. Lorch, R., Pugzles-Lorch and Klusewitz (1995) indicated the amount of typographical cuing (underlining) in a document will determine the recall effectiveness of the underlining, with more underlining resulting in less recall. Silvers and Kreiner's study (1997) revealed that students who read previously highlighted material showed little effect on performance versus non-highlighted material as long as the highlights were appropriate. Interestingly, if the highlighting was inappropriate, the study showed a notable decline in performance.

Recently, the effects of digital annotations have also been studied. Whether or not annotations aid cognition or benefit the reader, readers want to have the capability to highlight. While comparing personal annotations on paper versus shared annotations made on-line, Marshall and Brush (2002) noted the number of paper annotations far outnumbered the number of digital annotations (504 versus 98) with highlighting, underlining, or circling resulting in 82.1% of the paper annotations. The tool used for this experiment was WebAnn and the authors did not note any causes for the disparity in usage. Despite the disparity between students annotating on paper and annotating digitally, numerous contributors to the E-Book Functionality White Paper, DRAFT 1.0, January 2003 (Gibbons, Peters, & Bryan, 2003) have noted the desirability for annotation in digital documents.

2.4 Dimensions of Annotation

Annotation dimensions are important as they provide us with common categories for discussing annotations in their many forms. Marshall (1998) introduced seven dimensions of annotations. These dimensions are:

- *formal versus informal annotations,*
- *explicit versus tacit annotations,*
- *annotation as writing versus annotation as reading,*
- *hyperextensive versus extensive versus intensive annotation,*
- *permanent versus transient annotations,*
- *published versus private annotations,*
- *global versus institutional versus workgroup versus personal annotations*.

Formal annotations are usually restricted to filling out a form (in the digital world usually as a pop-up window) to allow metadata entry, while marginalia, highlighting etc., which are more free form, are considered *informal* annotations. Figure 2.1 illustrates examples of both informal and formal annotations. Formal annotations tend to be more distracting as the reader must now filter what information goes in what field.

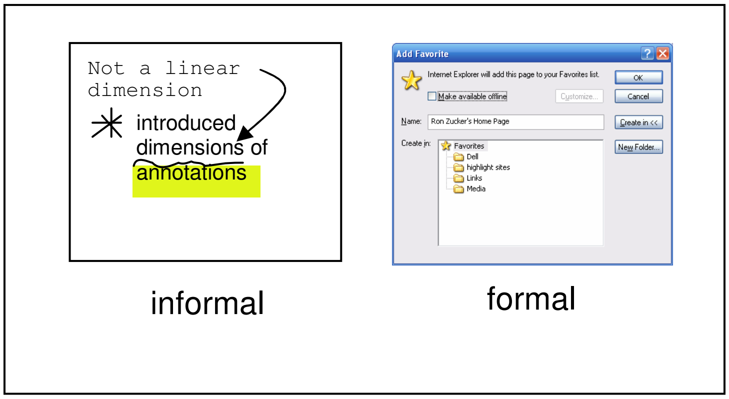


Figure 2.1 - Informal Versus Formal Annotations

Explicit annotations are usually complete sentences which contain the whole thought, usually intended for others to read, while *tacit* annotations are considered telegraphic and incomplete. Tacit annotation tends to be the most common form of annotation as the reader is doing only what is required, for example the asterisk, a highlighted word, or an arrow as illustrated in the informal example in Figure 2.1. HighBrow, the tool to be employed in this research, uses both explicit and tacit annotations. The highlighted context, because of its large size and content, may be thought of as explicit annotation and the highlighted keywords, which tend to be short and incomplete, as tacit annotation.

Annotation as *writing* defines the reader as a contributor to the original document. This particular dimension illustrates the fact that dimensions are not fixed since a reader may turn writer, especially with formal explicit annotations, then reader again.

Levy (1997) introduced a distinction between reading styles: *hyperextensive*, *extensive*, *intensive*. Marshall noted that this distinction may be extended to annotations: hyperextensive annotation, which involves link following and fragmentation over several sites; extensive annotation, Web scraping and/or summarizing from a number of sites; intensive annotation, the quantity of annotations within a given site.

The concept of *permanent* versus *transient* annotations is another dimension. While highlighting on paper may be considered permanent since it is difficult to remove the highlighted ink from paper, examples of transient paper annotations include Post-It®, inserted paper notes, or pencil markings. Digital highlighting is considered transient since the highlights can be easily hidden (and restored) as well as removed entirely.

Published versus *private* is a dimension that involves the intended audience of the annotation. Private annotations are never meant to be shared while public annotations are meant to be used in collaborative efforts. Marshall emphasizes the distinction between private and personal annotation since a book containing personal annotations can be loaned, given, or sold to others which contains personal yet not private annotations. Variations of the *personal* dimension include *global*, *institutional*, *workgroup*, and *personal* annotations. These variations have been the focus of many collaborative annotation tools (e.g., InterNote, Prep Editor).

2.5 Systems View of Annotation

Cousins, Baldonado, and Paepcke (2000) developed a systems view of annotations consisting of five components:

- *an annotation writing platform*,
- *an annotation reading platform*,
- *annotations*,

- *annotations correspondence*,
- *annotation target.*

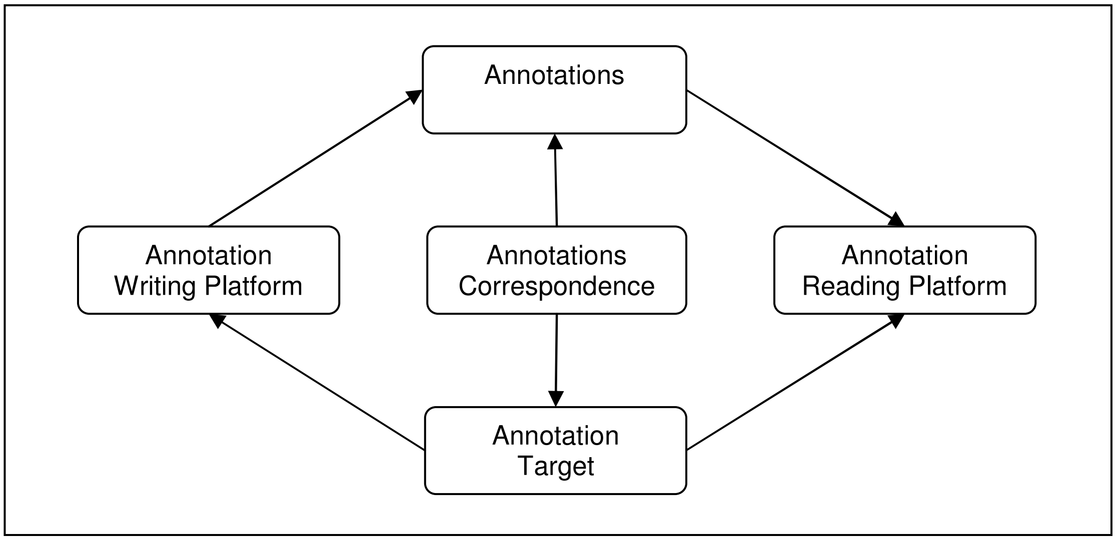


Figure 2.2 - The Five Annotation Components

Figure 2.2 shows the relationship between these five components. The arrows indicate that a relationship exists between the components. The annotation target is the source object being annotated and is present in both the reading and writing platform. The annotation writing platform creates the annotations, which are shown on the annotation reading platform and the annotations correspondence is responsible for maintaining communication between the annotation and the annotation target.

Three roles within the annotation system are defined as:

- The *author* of the original object that is being annotated.
- The *annotator* who comments on the object.
- The *reader* who is responsible for making sense of the commentary.

In this dissertation, the annotator is also called an active reader, the reader is called a passive reader, and the author is not addressed.

The system used to create the annotations is called the *annotation writing platform* which consists of: the input medium (real or digital *ink* and *keystrokes*); the location of the rendering platform for the annotation (which can be *collocated*, *separate*, or *both*); and the location (*margins*, *fixed overlay*, or *highlighted text*) and the capacity of the comment area (*limited* or *unlimited*).

The reader's interface when viewing the object and annotations is called the *annotation reading platform*. The annotation reading platform is broken into three parts: *display technique, hideability, searchability*. The display technique is how the reader can make a distinction between the original object and the annotation (e.g., *overlaid*, *inline*, *independent*). Hideability refers to the ability to display or hide the annotations from the reader. Searchability is having the ability to locate the annotation using a searching scheme. Possible choices for searchability are: *not searchable*, *limited search* (e.g., sequentially accessing annotations such as clicking next annotation), *searchable* (e.g., able to find an annotation by either entering in the annotation to be searched for or picking the annotation from a typically ordered list).

Annotation target refers to the object being annotated. In the digital world the object being annotated can be anything including text, images (fixed or moving), and sound. In this dissertation, the target is restricted to a Web document's text. The annotation target has three dimensions: *annotatable*, *physicality*, and *format dependence*. Annotatable has two values, *direct* and *indirect,* referring to whether the document itself is altered (direct) or are the annotations located elsewhere (indirect) without physically altering the original target. Physicality answers the question regarding the medium of the target. Is it *physical* (e.g., paper) or *digital* (e.g., appearing on a screen)?

Format dependence refers to the formats that the annotations can be applied to. For example, Portable Document Format (pdf) files can be viewed on the Web via plug-ins but many Web based annotation systems cannot annotate pdf files. The classifications for format independence are *strict*, *medium*, or *none*. Strict means a limited format (e.g., HyperText Markup Language (HTML) only), medium refers to multiple formats, while none works on all formats.

The fourth component is simply called *annotations*, which represent the characteristics of the annotations themselves. Annotations are divided into five characteristics: *scope, structure, published, liveness*, and *stature*. Scope refers to the intended audience of the particular annotation which can be *personal, workgroup, institutional,* or *global*. Cousins, Baldonado, and Paepcke (2000) point out, in some applications, annotators may vary the scope of the annotations intending for some annotation to be personal, others to be workgroup, etc. Published refers to the intended audience and availability of the annotations. If the intended audience of the annotation is the annotator, then the annotation is considered unpublished. If the annotation is intended for a wider audience then the annotations are published. The two characteristics of scope and published appear to be the same but distinctions may be made. For example, a personal annotation may be published, however the annotation was made for an individual's own use, as in a comment written in a textbook, then loaned or resold making the personal annotation gain a wider audience. Structure is the characteristic that describes the manner in which annotations are created. A scribble on paper is an example of an *informal* annotation, while a form (metadata entry) with predetermined fields is an example of a *formal* annotation.

The final component, *annotations correspondence*, describes how communication between the target object and annotation is specified and maintained. Annotations correspondence has four characteristics: *mutability, anchoring granularity, anchoring technology,* and *robustness*.

Mutability defines how the annotation is maintained if the underlying data changes. With paper documents, the text is considered *immutable*, that is, it does not change. This immutability is true even if a newer edition of a text document is produced, since the original document remains present in its current state. In the digital world, the document at a given location may, and often does, change. Will the annotation move if the text moves? What happens when the annotated text is changed or deleted? If the annotations satisfactorily answer these questions, the annotations are considered mutable.

The level within the document that the annotation target can reference is called anchoring granularity which may be: *page* (as in Internet Explorer's "Favorites"), *character*, or *pixel* (for multimedia objects). Anchoring technology refers to how the annotation is located in the document, either by *juxtaposition* or by *Uniform Resource Locator* (*URL)*. Juxtaposition means anchoring the annotation to a location within the document, usually near the object that is being annotated, while URL means annotations are summarized or retrieved at the document level.

More and more Web documents are becoming dynamic or mutable. Robustness is the characteristic that describes "the ability to modify correspondence without affecting annotations". Robustness may be identified as *removable*, *URL-stable*, or *permanent*.

Table 2.2 lists several annotation systems with a breakdown of each of the five system components. Since all properties are not discrete, the properties have been marked with plus (+) or minus (-) to show the degree of fit with + meaning exceeds the requirement and – indicating to a lesser degree. The asterisk (*) in the table indicates that the property exists with caveats.

2.6 Highlighting Browsers

Highlighting is the most common form of annotation (Brennan, Winograd, Bridge, & Hiebert, 1986; Marshall & Brush, 2002). Highlighting has been used for annotating using marker and paper, with software word processors, such as WordPerfect and Microsoft Word, and highlighting browsers.

The concept of a highlighting browser is not new. The roots for annotating browsers began in 1945 when Vannevar Bush envisioned a device for interactive information annotation and path tracking which he called ”memex” (short for MEMory EXtender). While the technical description is quite different from current World Wide Web implementations (memex was based on microfilm technology), the concept of a retrievable and annotatable digital document is being realized.

Theodor Nelson, a Sociology instructor at Vassar College is credited with introducing the term hypertext in 1965 in his talk “Computers, Creativity, and the Nature of the Written Word.” Nelson offered the challenge “to design “hyperfiles” and write “hypertext” that may have more teaching power than anything that could be ever printed on paper” (*XANADU® ARCHIVE PAGE.*). Nelson invented the P.R.I.D.E. (Personalized Retrieval Indexing and Documentary Evolution) system, designed to translate passages of

Table 2.2 - Comparison of Annotation Systems (Cousins et al., 2000)

	Writing Platform *1. input medium* *2. platform* *3. size*	Reading Platform *1. display technique* *2. hideability* *3. searchability*	Annotation Targets *1. annotatable* *2. physicality* *3. format dependence*	Annotations *1. published* *2. scope* *3. structure* *4. liveness* *5. stature*	Annotation Correspondence *1. granularity* *2. mutability* *3. technology* *4. robustness*
Direct paper	1. ink 2. collocated 3. margins	1. overlaid 2. not hideable 3. not searchable	1. direct 2. physical 3. medium	1. unpublished 2. personal 3. informal 4. inactive 5. fragment	1. pixel 2. immutable 3. juxtaposition 4. permanent
Post-Its™	1. ink 2. separate 3. fixed overlay	1. overlaid 2. hideable 3. not searchable	1. direct 2. physical 3. none	1. unpublished 2. personal 3. informal 4. inactive 5. document	1. page+ (*) 2. mutable 3. juxtaposition 4. removable
Annotated Edition	1. keystrokes 2. depends (*) 3. unlimited	1. inline 2. not hideable 3. limited search (*)	1. direct 2. physical 3. strict	1. published 2. global 3. informal 4. inactive 5. document	1. page 2. immutable 3. juxtaposition 4. permanent
X Libris	1. ink 2. collocated 3. margins	1. overlaid 2. hideable 3. limited search	1. direct 2. digital 3. strict	1. unpublished 2. personal 3. informal 4. active 5. fragment	1. pixel 2. immutable 3. juxtaposition 4. permanent
MVD	1. ink- (*) 2. collocated 3. margins	1. overlaid 2. hideable 3. limited search	1. indirect 2. digital 3. none	1. published 2. global 3. informal 4. active 5. document- (*)	1. pixel 2. mutable- (*) 3. juxtaposition 4. URL-stable
MS Word	1. keystrokes 2. collocated 3. unlimited	1. inline 2. hideable 3. searchable	1. direct 2. digital 3. strict	1. unpublished 2. personal 3. informal+ (*) 4. active 5. fragment	1. character 2. mutable 3. juxtaposition 4. removable
ComMentor	1. keystrokes 2. collocated 3. fixed overlay+ (*)	1. independent 2. hideable 3. searchable	1. indirect 2. digital 3. medium	1. published 2. varies (*) 3. informal 4. inactive 5. document	1. character 2. mutable 3. URL 4. URL-stable
NotePals	1. ink 2. separate 3. unlimited	1. independent 2. hideable 3. limited search	1. indirect 2. varies (*) 3. none	1. published 2. workgroup 3. informal 4. inactive 5. document	1. document+ (*) 2. immutable 3. URL+ (*) 4. removable
Tapestry	1. keystrokes 2. collocated 3. unlimited	1. independent 2. hideable 3. searchable	1. indirect 2. digital 3. strict	1. published 2. workgroup+ (*) 3. formal 4. active 5. fragment	1. document 2. immutable 3. URL- (*) 4. permanent- (*)
Notable	1. keystrokes-(*) 2. separate 3. unlimited	1. independent 2. hideable 3. searchable	1. varies (*) 2. varies (*) 3. none	1. published 2. varies (*) 3. varies (*) 4. inactive 5. document	1. document 2. mutable- (*) 3. URL 4. URL-stable

material into machine language and filed in the machine in any sequence. The writer could later retrieve the information in any sequence, freeing him from memorizing the ideas (*XANADU® ARCHIVE PAGE.*).

In the mid 1990s, with the growth of the World Wide Web, interest in annotating Web documents began in earnest. These attempts were modeled after existing document processing software, using techniques such as Post-it® notes, highlighting, underlining, etc. Implementations such as Amaya, Annotea, iMarkup Client, Xlibris, YAWAS, i-Lighter, and others provide Web annotation and collaboration with mixed results and acceptance. The existing Web annotation software can be described using summarization capabilities, user interface, and type of annotations supported.

Annotea (Annotea project, 2005), which is the annotating portion of the Amaya project (Vatton, 2007), was the first experimental implementation of Web annotations based on the Resource Description Framework (RDF) (Herman, Swick, & Brickley, 2007). The annotations are anchored at the document level and require typing by the reader. Figure 2.3 is a screen shot of a sample annotation.

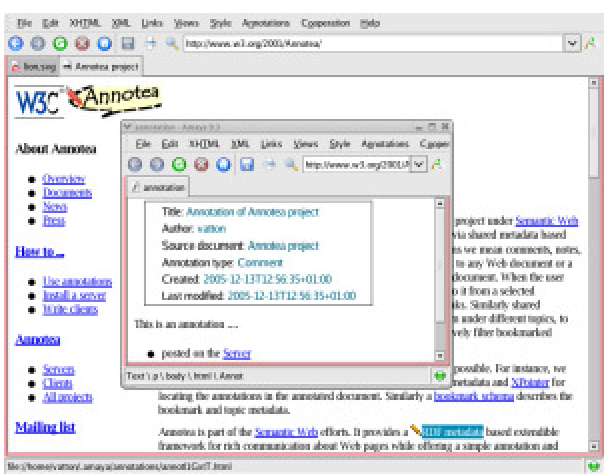


Figure 2.3 - Screen Shot of an Existing Annotation Using Amaya (Vatton, 2006)

Piggy Bank, an extension of FireFox, part of MIT's Simile (Semantic Interoperability of Metadata and Information in unLike Environments) project, may be considered a hyperextensive annotation tool, since it provides a way of using extracted content via third party screen scrapers from multiple sites and combining them into a single document. The user can collect, save, search, browse, share, and retrieve document information based on numerous properties (called facets) of the document. When content is scraped from a visited Web site and saved, PiggyBank creates an index of the content for future searching using the Semantic Web (Herman, 2007) and RDF.

As Quan and Karger (2004) noted, citing the classic end-to-end argument (Saltzer, Reed, & Clark, 1984), the consumer of information is the best judge of what is important and how to use it. Piggy Bank permits the user to add personal tags to the scraped content by typing in keywords via a pop-up form (Huynh, Mazzocchi, & Lee, 2007). The user can later search for information based on these tags.

iMarkup Client is a commercial product that permits annotation using sticky notes, free form drawings, text markups, and voice. In addition to HTML files, iMarkup Client also supports pdf files. Text markups include underlining, highlighting, bolding, and italicizing. The user has the capability of typing in (or even speaking) annotations to the text markups and categorizing the text markups. Searches must begin at the URL level, either by Web site or page, creation, viewed, or modified date, or by the author's name. Within a URL the search can be expanded to text markup, sticky notes, or paint brush (free form drawings using the mouse). The interface is quite simple but could cause an interruption in the reading activity because of the various menus required to complete searchable markup.

XLibris is an approach to annotation based on the paper document metaphor that utilizes unique hardware and software. Xlibris uses a pen tablet display that emulates the appearance of a sheet of paper which can be attached to a conventional computer or standalone, portable pen computers (Schilit, Golovchinsky, & Price, 1998). Schilit, Golovchinsky, and Price (1998) point out, “annotating with a pen requires little cognitive overhead compared to typing or to selecting text with a mouse and issuing a command.”

The XLibris system also adds many features beyond the paper document metaphor including: *query-mediated links* (XLibris will search for related articles based on the annotations made by the reader) and “*The Reader’s Notebook*”, which is a summary of the annotations made for the given document or all documents. To help the reader understand the summary in The Reader’s Notebook, each annotation also includes the document title and page number to help interpret the annotation.

YAWAS (Yet Another Web Annotation System) is a simple text highlighting annotation system (Denoue, 2005). YAWAS is a JavaScript plug-in for Internet Explorer. The user simply selects text with the mouse and right-clicks to add an annotation. The user may also add a note concerning the highlighted material. The highlighted material may be retrieved using YawasQuickSearch. YAWAS offers the advantage of creating bookmarks or favorites at the word level rather than the document level as found in most annotation software.

In evaluating the usage patterns of YAWAS in 1999, Denoue noted that the users were more apt to highlight then to provide annotated descriptions. In 2005, Denoue has maintained the annotation portion of YAWAS, but is now concentrating more on the performance and storage location of the highlighting data. Denoue was initially concerned about privacy, storing annotations on the client side. Denoue has now added centralized storage in order to provide portability and collaboration.

i-Lighter (*i-Lighter The Yellow Marker on the Web.* 2006) is a commercial product similar to YAWAS. i-Lighter is a plug-in for Internet Explorer and FireFox that allows the user to highlight text and graphics from Web documents and add notes similar to Post-its®, called i-Notes. Highlighted content may be retrieved by opening a retrieval page that contains a document viewer showing the highlighted document, a directory tree of previously highlighted pages, and a list of the highlighted documents within the selected directory.

All of the aforementioned annotation systems offer powerful tools for annotation, summarizing, and searching. The existing systems require typed in interpretive notes, called "tags", if the user wishes to explain the annotations. Switching from selection (using a mouse or a digital pen) to typing creates a disruption in cognition.

Another drawback to these systems is the compromise between importance and interpretation. Unless there are typed annotations to aid interpretation, the user must return to the original source document to help interpret the annotations. To avoid returning to the original source, the user is confronted with the option of summarizing more content at the risk of losing importance.

2.7 Existing Architecture

Figure 2.4 depicts a general architectural approach, called ComMentor (Roscheisen, Mogensen, & Winograd, 1997), for highlighting Web documents. The user requests a document from a document server; the document is then brought into a document synthesizer and combined with data arriving from meta-information servers.

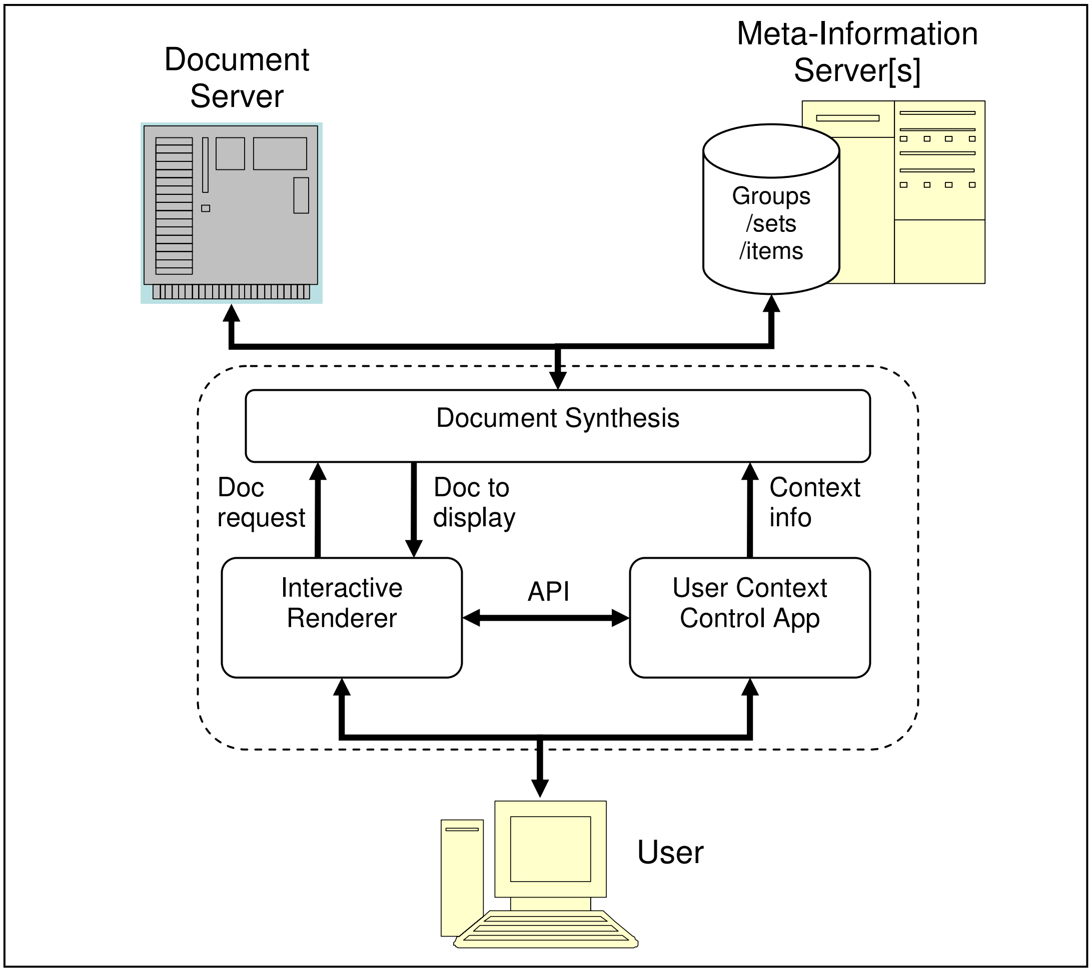


Figure 2.4 - General Platform for Providing Annotation Via Third-Party Value-Added Information Providers (Roscheisen, Mogensen, & Winograd, 1997)

Meta-information servers supply the annotation data and in the case of collaborative annotations, group and user information. The user context control application is responsible for filtering the annotations that the user is allowed or wants to see. The document synthesizer then inserts the qualified annotations into the local version of the document where they are rendered onto the screen.

This architecture has the advantage of being implemented as a plug-in for existing browsers allowing users to use the browser of their choice (providing the browser accepts plug-ins) as the interactive renderer.

2.8 Summary

This chapter defined annotation and the multiple forms of annotation with the reasons for each type of annotation. Next, the cognitive effects of annotation were explored. The latter portion of this chapter concentrated on digital annotation. Annotation has matured to a point where dimensions and a systems view have been established which "helps to compare systems and to design new ones" (Cousins, Baldonado, & Paepcke, 2000). In addition, this chapter provides an overview of existing work with Web based annotation and more specifically highlighting text on the Web. Included is a general architectural approach to Web based annotations.

Marshall (1997) has pointed out short highlighting aids memory and longer textual notations aid interpretation. Marshall asserts the act of switching from highlight marker to pen can be distracting, thus students who highlight write fewer marginal notes than persons who underline with pens. Further research by Marshall and Brush (2002) has shown approximately 82% of the people who made personal annotations use highlighting, underlining or circling while only 4.6% used notes only and 7.5% used a combination of notes and highlighting, underlining or circling. Highlighting aids memory but is not good for interpretation; textual annotation is good for interpretation but does not aid memory.

Can we have it both ways with a single annotation scheme? Context highlighting is a single annotation scheme which allows the annotator to note keywords through highlighting and add interpretive text by highlighting the supporting text surrounding the keywords. The annotator is not required to switch from mouse to keyboard and back and, since the context is already present, there is no need to type it in. The same highlighting action is used for both! This may lead to more awareness of the important portions of the document and a much better summary that may be of use to both annotator and reader.

HighBrow is a prototype Web browser that enables both keyword and context highlighting. HighBrow is introduced in the following chapter, and compared with the dimensions, the systems view and the architecture discussed in this chapter.

Chapter Three:

HighBrow

This chapter introduces HighBrow, the context enabled highlighting browser and its many versions and sub-versions. While many applications were developed and used during the experiments, HighBrow is the key tool that was developed for context enabled highlighting browsing.

3.1 HighBrow, A Context Enabled Highlighting Browser

There are a number of annotation systems in existence that support highlighting Web pages; however, context highlighting is new. A decision had to be made to take an existing system and add context highlighting or begin from scratch. Modifying an existing system presented many obstacles including access to source code, copyright issues, etc. In addition, the method for determining where to physically locate the highlight in the existing systems was not deemed satisfactory, since a search was required to determine where to place the highlight. Another consideration was the need for complete control of the experiment by preventing the participants from surfing to other Web pages or bypassing the process of highlighting altogether.

In order to facilitate the development of a context highlighting browser, create a more efficient method of highlighting, and observe and collect data in a controlled environment, the decision was made to create a context highlighting browser from scratch. The new browser was named HighBrow, which is a contraction of HIGHlighting

BROWser. The following sections describe HighBrow's architecture, interface, process, dimensions, systems view, and testing scheme.

3.1.1 HighBrow Architecture

The architectural goal for HighBrow is to ensure accuracy, improve speed, and simplify the process. This section provides an overview of the architecture of HighBrow and shows how HighBrow's architecture differs from the ComMentor architecture, described in Chapter Two. Figure 3.1 provides an overview of the HighBrow architectural components. The document server is any server that provides HTML content. Similar to the ComMentor architecture, the request and document are provided using standard Hypertext Transfer Protocol (http) and the source document is never altered. The meta-information is stored on a database server which may be any server capable of handling a relational database.

Rather than using a Resource Description Framework / Extensible Markup Lanuague (RDF/XML) model (Annotea project. 2005), relational database tables are used with filtering being provided via SQL select statements. HighBrow is located on the client's computer and is responsible for requesting the Web document from the document server and any corresponding highlight data from the database server. HighBrow is also responsible for detecting and storing any new highlights, modifications to existing highlights, and/or removing highlights that have been deleted.

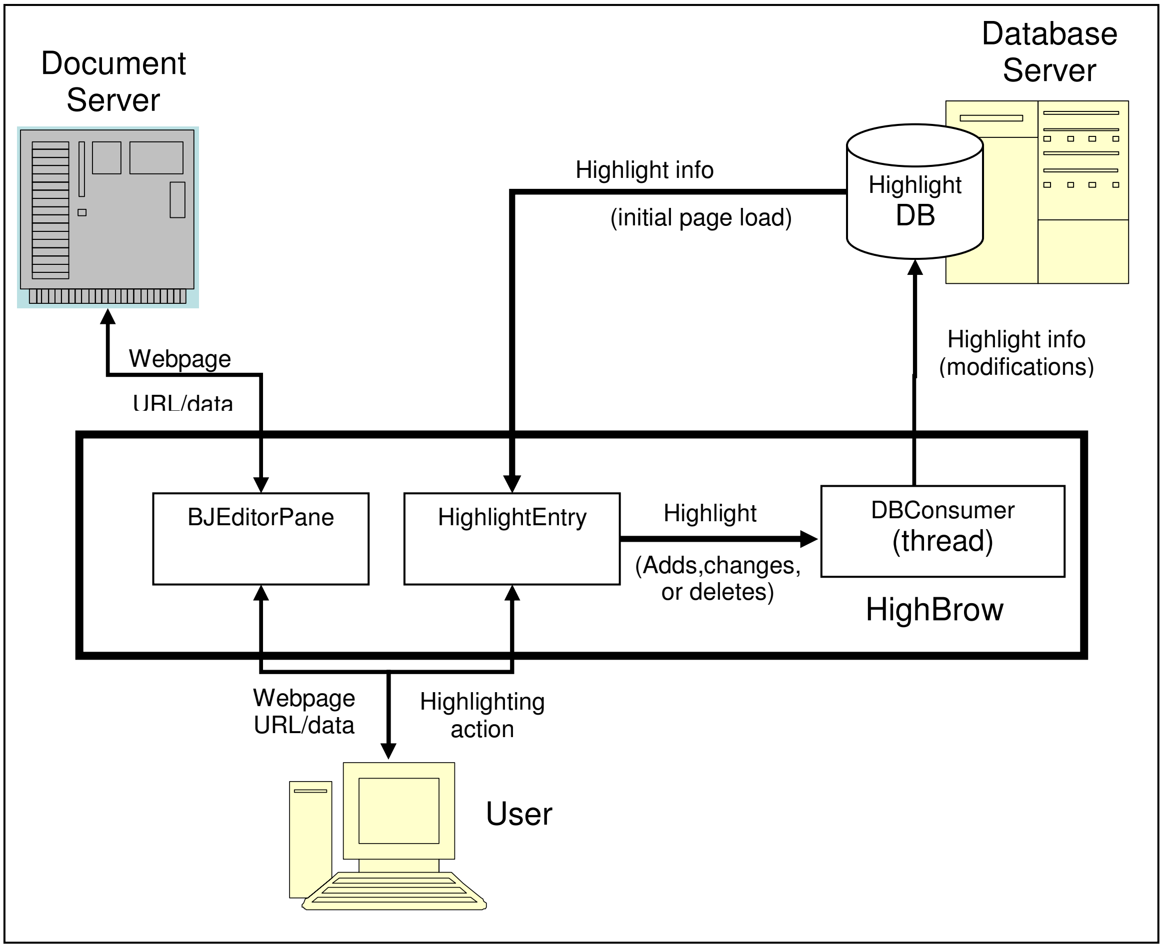


Figure 3.1 - HighBrow Architectural Components

In order to improve performance, HighBrow caches the highlights for the current pages and uses threads to make updates to the database. To reduce database overhead, the DBConsumer class provides a queue for requests which resides in the thread. As long as there are active requests in the queue, the thread is allowed to live and the connection to the database is maintained. Once the queue is emptied, the connection is closed and the thread dies.

HighlightEntry is a Java class used to form the local cache for highlight metadata. HighLightEntrys may be found in two lists: one list for the current page; the other list for highlights on other pages. Only the current page list, along with the database, is updated during highlighting since the highlights on other pages are not affected by the current page modifications. When a new page is requested, both lists are updated, the keywords are loaded as before, but the keywords from the page just visited are moved into the keywords other pages list. HighBrow does not permanently store highlighting data on the client machine.

A user may request a page either by typing in a URL address in the address field, by clicking on a link in the document, or selecting the “Go to Webpage” menu item within the context summary page. After the Web page is located, it is loaded into the JEditorPane on HighBrow and a separate request is made for all highlight metadata related to that particular address for that particular user.

Because the highlighting process is dynamic, and highlights can easily change state from context to keyword and back, the task of determining which entry was keyword or which entry was context was left to HighBrow rather than encoding the metadata. HighBrow determines whether the highlight is a keyword or context by determining the start and end locations relative to other highlights. The highlighting data are retrieved using the userid, address, and starting and ending locations of the highlights. The retrieved data are sorted with the starting location descending and the ending location ascending, which moves keywords ahead of context. Because it is possible to have a keyword without context, but impossible to have context without at least one keyword, the order of retrieval makes the process of determining whether the highlighted passage is keyword or context relatively simple. HighBrow marks everything as a keyword unless the boundaries surround the previous entry which means that entry is a context entry.

The database server uses a relational database to store the tables required for the experiment. The schema includes three tables: password, reference, and history, which are described in detail in Table 3.1, Table 3.2, and Table 3.3. The italicized items in each table indicate the primary key. The history table is used only to gather data about HighBrow usage for the experiments and is not required and should be eliminated in an actual implementation.

Table 3.1 - Password Table Description

Attribute name	Format	Description
USRID	VARCHAR2(8)	Userid
SEX	VARCHAR2(1)	Gender of user (M/F)
AGE	NUMBER(38)	Age group of user
HIGHLIGHT	NUMBER(38)	Highlighting habit of user
PRINT	NUMBER(38)	Printing habit of user
COURSE_NAME	VARCHAR2(50)	Initially meant to track users by course name now groupid

Table 3.2 - Reference Table Description

Attribute name	Format	Description
URID	VARCHAR2(8)	User's userid.
GRID	VARCHAR2(8)	Group Id
ADDRESS	VARCHAR2(150)	URL address
VIEWED	DATE	Date and time of event
POSSTART	NUMBER(38)	Starting position for highlight
POSEND	NUMBER(38)	Ending position for highlight
QUOTE	VARCHAR2(4000)	Text of Highlight
KEYWORD	NUMBER(38)	

Table 3.3 - History Table Description

Attribute name	Format	Description
URID	VARCHAR2(8)	Browser Userid
ADDRESS	VARCHAR2(150)	URL address
EVENT_TIME	DATE	Date and time of event
EVENT_TYPE	NUMBER(38)	See HighBrow History events table for codes used
POSSTART	NUMBER(38)	Starting position for highlighting or deleting highlighting
POSEND	NUMBER(38)	Ending position for highlighting or deleting highlighting
SESSION_ID	DATE	A unique time code to identify a session

Table 3.4 shows the event types and descriptions of the events captured by HighBrow. This data were invaluable for recording annotating behavior during the experiments.

Most existing annotation software utilize plug-ins and/or RDF/XML technology to alter the local copy of the page in order for it to be rendered by the browser. In contrast, HighBrow uses the highlighting capability provided by the JEditorPane class supplied by Java, which involves overlays. This has advantages and disadvantages.

Table 3.4 - History Event Types

Event Type	Event Description
10	login
20	URL begin from select
30	URL begin Summary ref other page
40	URL begin
50	highlight enter
55	highlight update
60	highlight delete
70	Go to highlight ref this page
80	Show context from keyword summary
85	Show context summary for current page
90	Print context
95	Copy Context to file
100	Close context
110	Print document
120	hide highlights
130	show highlights
140	delete all highlights
150	URL end
160	logout

The advantages include a fast, accurate, and relatively easy way to highlight text. Earlier annotation systems attempt to create a separate document, add tags, and, if necessary, modify internal links to point to the new document (Yee, 2002). Other systems use an intermediary to redevelop the original document, attaching the annotations (called *anchoring*), and then redisplaying the new document. In addition to being a bit more time consuming, this process has no way of being 100% certain of highlighting the same passage the reader intended to be highlighted. Due to the browser interpreting and filtering markup tags, the location of the displayed text does not necessarily agree with the original document (Denoue & Vignollet, 2000), the location is approximated and then a search is introduced. If the highlighted text is ambiguous then it is possible for the incorrect text to be highlighted.

Document Object Model (DOM) level 2 introduced range objects which are subparts of the original document facilitating the anchoring of the highlight with the original document (Denoue & Vignollet, 2000).

Highlighting within the JEditorPane uses the already formatted text, thus rendering the highlight is fast and accurate. Since the highlight location is the same as the start and end location of selected text, there is no search requirement, HighBrow simply adds a Java highlight to the text at the given location.

A disadvantage of the JEditorPane is that the highlighting features are quasi proprietary, making their use restricted to the Java platform. Another disadvantage is, by not being a plug-in for other browsers, the entire browser with all of its features must be developed from scratch.

3.1.2 HighBrow Interface

The actual interface for HighBrow supports the following Human Computer Interface (HCI) concepts: transfer, proximity, visibility, and feedback.

The overall layout of the browser provides moderate positive transfer from existing browsers as represented in Figure 3.2. The document title, file menu, elementary navigation items, and the window buttons (minimize, maximize/restore down, and window close) are located in generally the same location. The keyword indices are located in the same area as the history panel would be, if opened. The operations available on HighBrow are much the same as existing browsers allowing the user to: navigate, by clicking on a link or entering a URL address; highlight, by dragging the mouse over text; activate the mouse by right clicking and showing a menu of actions. The usual choices when the right mouse is clicked, for example copy, paste, have been dropped since creating a highlight is similar to copy and paste. If the user chooses to copy or paste, these options are still available via the key sequences (control-c and control-v).

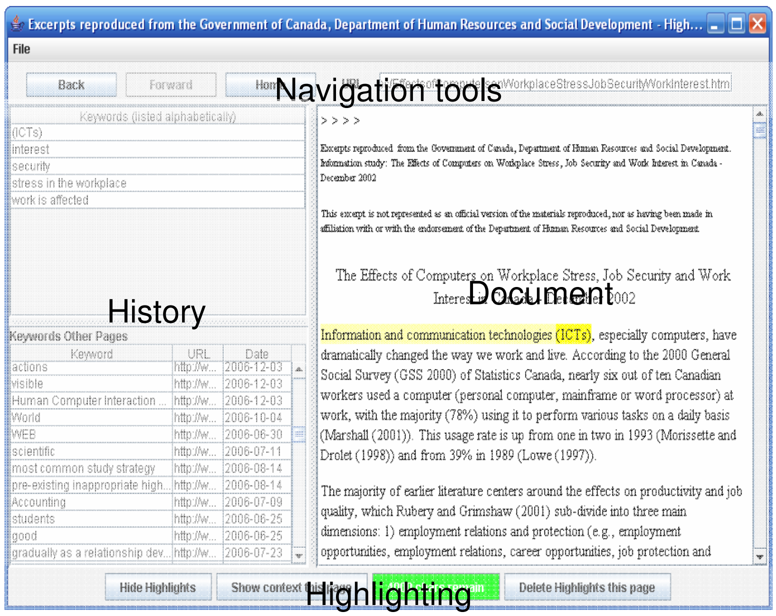


Figure 3.2 - Browser Layout

The gestalt principle of proximity groups is used for document navigation, keyword indices, editor pane and highlighting features to facilitate ease of use and learning.

Document navigation components are placed along the top, historical data concerning highlights are placed in a split pane along the left (similar to the way the history is presented using Internet Explorer) and the content pane is located in the right center. Since HighBrow allows highlighting, highlighting controls are visible along the bottom of the browser.

The navigation controls include the familiar back, next, and home keys as well as the address box that can be used to enter a URL address by typing it in or to display the current Web page's address. The user is also able to navigate by clicking on a hypertext link in a manner similar to other browsers.

Unlike traditional browsers, a resizable history frame is always present on the left side and this frame contains two resizable frames stacked vertically. The top frame is a selectable table of keywords organized alphabetically. Clicking on any of these keywords will position the content frame so that the beginning of the context surrounding that particular keyword is brought to the top of the pane. If the keyword/context is at the bottom of the document, then the keyword or context will not be shown at the top but the page will be shown within the context pane. This occurs because it is impossible to scroll beyond the end of the document.

The bottom frame contains a table of keywords located on Web pages other than the one currently being shown. This list is functionally similar to the favorites list found in existing browsers, except it can be sorted in alphabetical order by keyword or URL address. The list can also be sorted by the date that the highlight was made in descending order so that the most recent entries are on top. Clicking on these entries will open up a separate window containing a summary of keywords and the context surrounding them. If the user wishes to view the original document, the context summary has a menu selection that allows the user to navigate the content pane to the appropriate site.

The content pane contains the Web document and presents the data in fashion similar to other browsers. The major difference between HighBrow's content pane and traditional browsers' content panes is the ability to highlight and retain keywords and context.

3.1.3 The Highlighting Process

Users have shown a desire for simplicity to facilitate annotation (Obendorf, 2003) and lightweight annotation functions (Fu et al., 2005). The highlighting process was developed with simplicity in mind, turning off, or removing entirely, options if they are not viable. This section describes the process of highlighting keywords and context.

When a user drags the mouse over content within the content pane, the text is highlighted in light blue as shown in Figure 3.3.

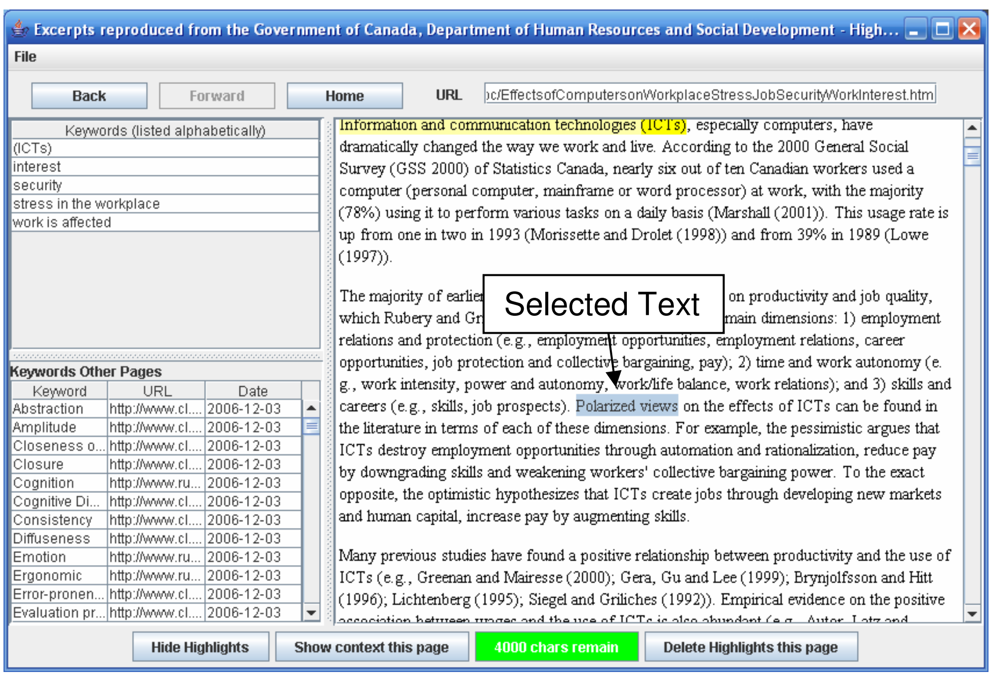


Figure 3.3 - Example Text Selection

If the user then right clicks the mouse, either one or two actions occur. If the data have not been previously highlighted, the text previously highlighted in light blue is highlighted in bright yellow and the words highlighted appear in the keywords list as shown in Figure 3.4.

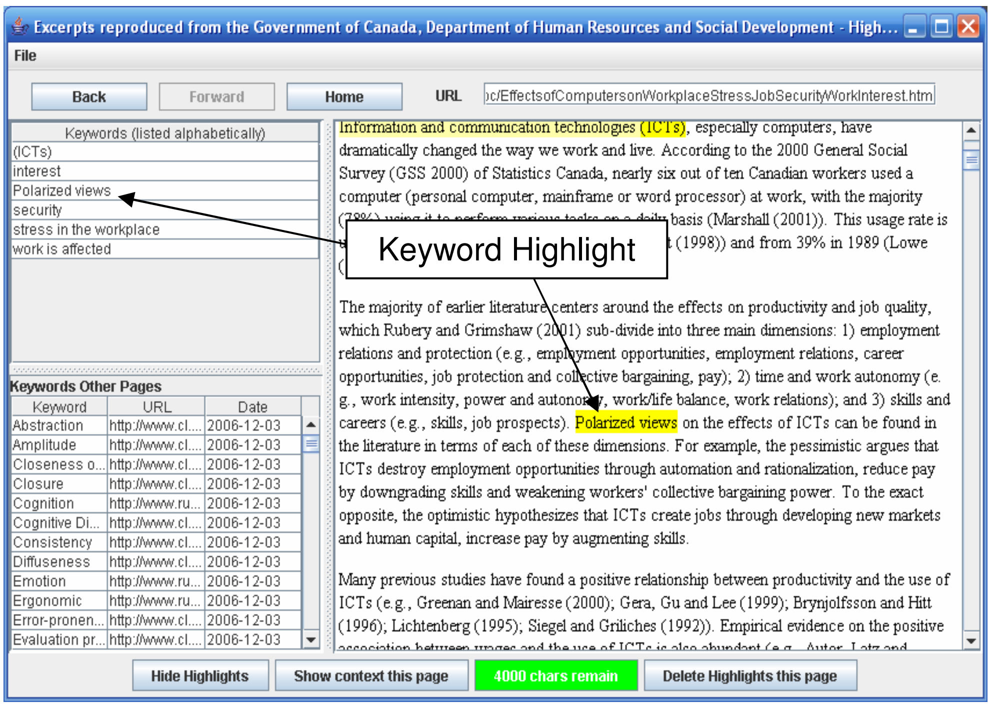


Figure 3.4 - Keyword Highlight

If the words highlighted include previously highlighted material the action is significantly different. Highlighting previously highlighted text can mean several different things. The user may wish to:

- modify the size of the existing highlight,
- add a keyword to an existing context,
- add context to an existing keyword,

- delete the context highlight,
- delete the keyword highlight

The choices are reduced by HighBrow, allowing the user to only perform the actions that make sense for that particular situation. Figure 3.5 is an example of a possible pop-up window when previously highlighted material is included in the selection. The pop-up window is placed to the immediate right and below (as room allows) the highlighted material, thus allowing visibility of the highlighted material. Placing the pop-up adjacent to the current highlight reduces the distance the user must move the mouse from the highlight to the menu choice. If a toolbar pull down menu or a fixed button were used the distance could be far greater. The pop-up window takes advantage of a concept known as Fitt's Law, which addresses the amount of time a user takes to reach a target, given by the equation:

$$\text{Time} = K \log_2(A/W+1)$$

where K is a constant, A is the amplitude (distance between the start and target), and W is the width of the target being sought. In addition to decreasing the time to make the selection, the location of the pop-up reduces interruptions to the task at hand, which is reading and highlighting.

Figure 3.5 - Example Pop-Up Window

If the user highlights text that includes a previously highlighted keyword, the choices will be to resize the keyword, add context, or delete the highlight. Since context has not yet been established, it does not make sense to adjust context size.

If the user highlights text within existing context, the choices will be to resize the context, add a keyword, or delete the highlight. Since context is established, it does not make sense to add a context.

Since a user may begin and end highlighting at the character level, the software must make decisions with respect to the extent of highlighting. For example, if the user wishes to highlight context and stops midway into an existing keyword and the user selects add context, then the program will extend the context to include the entire keyword since the keyword is considered part of the context.

Figure 3.6 is a screen capture of a document with the context highlighted and the pop-up window activated.

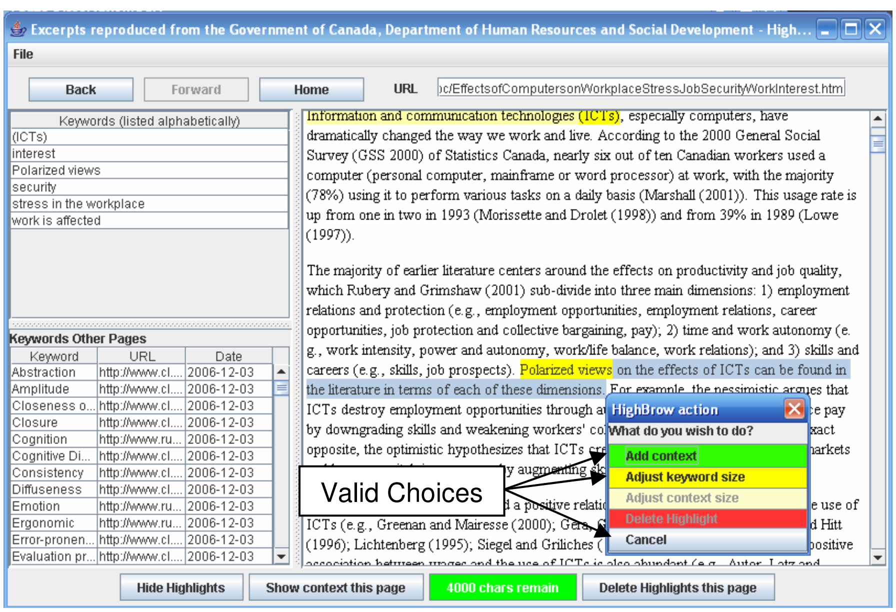


Figure 3.6 - Context Highlight with Pop-Up

If the user selects add context, the final results are shown in Figure 3.7. The context is displayed in pale yellow. Note that the context is not added to the keyword list.

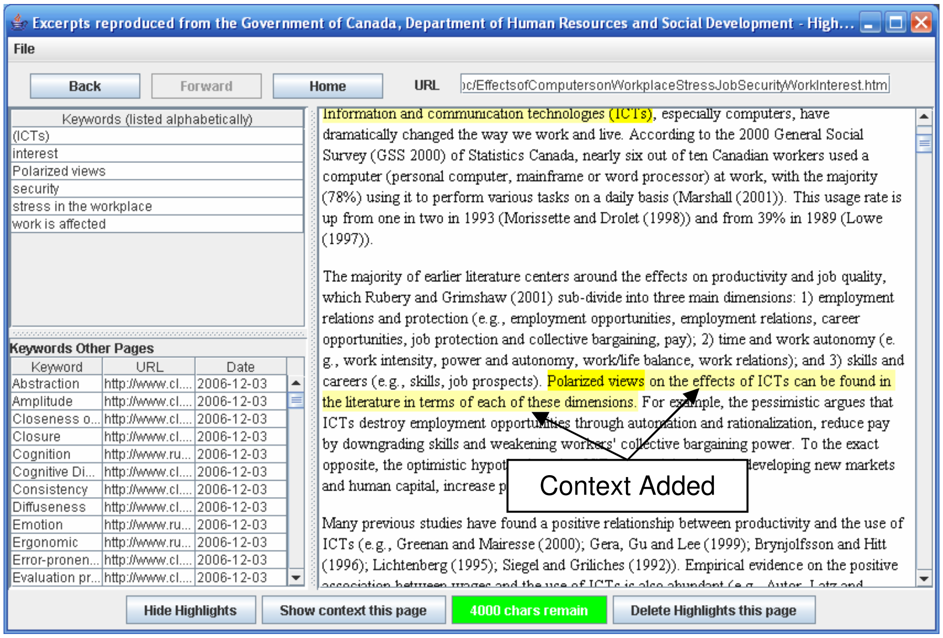


Figure 3.7 - Added Context Highlighting

If the user inadvertently right clicks the mouse and nothing has been selected then the pop-up warning in Figure 3.8 is issued. The upper left hand corner of the pop-up is placed where the mouse was clicked to reduce the distance of travel for the mouse.

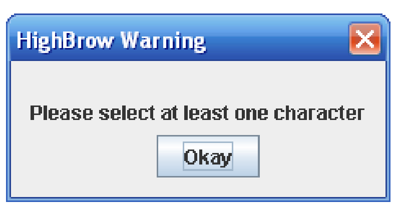


Figure 3.8 - Pop-Up Warning Window

3.1.4 HighBrow Context Summary

One of the advantages of digital highlighting is the ease of organizing the highlighted text. In addition to providing the index for finding the keywords, HighBrow is able to produce a summary consisting of keywords and the surrounding context. Figure 3.9 is an example context summary. The context summary continues to show the keywords in bright yellow, but the context is no longer in color as all remaining displayed text is, by definition, context.

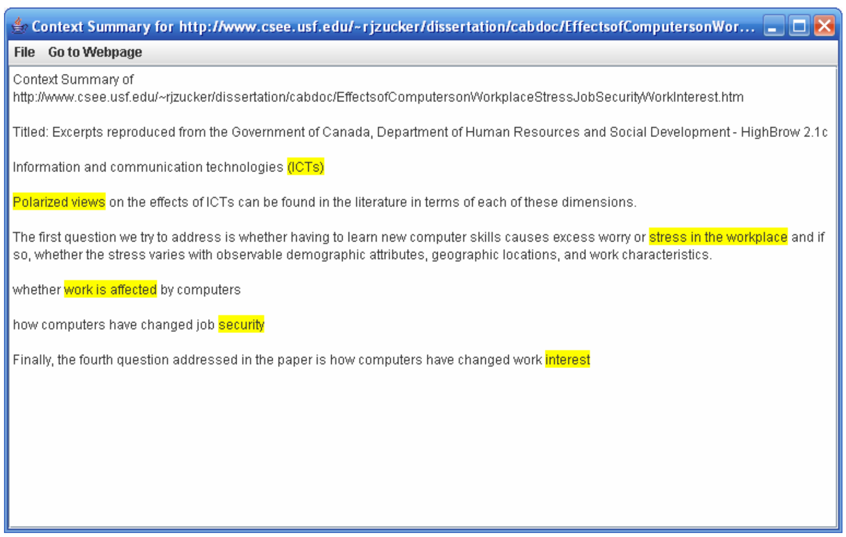


Figure 3.9 - Document Summary

Every context entry is separated by a double space, even if the highlighted contexts appear on the same line, to delimit each context selection. If the user wished the entries to appear on the same line, then the user would have simply made a single larger context.

The “File” pull-down menu allows the user to print (with highlights) or copy (without highlights) the summary to a file. The “Go to Webpage” is for navigating to the original Web page for this particular summary. Clicking on this button is necessary only if the user selects from the “Keywords other Pages” links, since the document pane does not change when the context summary is opened.

3.1.5 HighBrow Dimensions

The dimensions of the annotations (Marshall, 1998) with respect to HighBrow are described in this section. HighBrow restricts the annotation type to highlighting only.

To avoid any distractions while reading, the highlighting in HighBrow is considered *informal*. This distinction is being made because some applications make/allow the user to fill out a brief form regarding the highlighted text. In HighBrow, the biggest interruption in the reading process is selecting from a brief menu. No forms exist that require the user to shift input from the mouse to the keyboard.

The highlighting in HighBrow is both *explicit* and *tacit*. Context highlighting is considered to be explicit and eliminates the need for most marginalia or added text. Keyword highlighting is considered tacit since their meaning is not expressed directly since the keywords make little or no sense when taken alone.

In the sense of contributing to the original text, the reader using HighBrow never becomes a true *writer*. The reader may change roles slightly as the reader considers the amount of context to include in support of the keyword or phrase. In this way, the reader can be considered a writer without the interruption of physically writing or typing.

With respect to *hyperextensive, extensive,* and *intensive*, the highlighting in HighBrow can be considered to be hyperextensive and intensive. HighBrow is hyperextensive because every highlighted keyword or phrase immediately becomes a link, both within the document and when viewing other documents. HighBrow is also intensive as there is virtually an unlimited amount (based solely on the amount of storage allowed by the database) of highlights allowed per document. There is a limitation on the size of any one highlight of 4,000 characters, imposed by the database, however, this should not be considered limiting as the reader may highlight multiple 4,000 character segments.

Despite the fact that HighBrow supports Web scraping, it is not considered extensive, as in its current state HighBrow does not easily allow combining summaries from multiple sites. HighBrow allows the user to copy the summary to a file but combining the copied files is not one of the objectives or features of HighBrow.

As in most software implementations of annotations, the highlights in HighBrow are considered *transient*. The highlights can be hidden, redisplayed, or deleted permanently. The deletion process can be by keyword, context, or document.

In its current state, the highlights created by HighBrow are considered *private*. The highlights are only shown to the individual user. Earlier versions of HighBrow did allow for *published* highlighting, showing three different levels (public, group, and individual) in different colors and was able to dynamically filter at the click of a button. The different published levels were removed because the experiments dealt with individual performance and were not group related, however the database field *grid,* short for group id, has been retained and is used for analysis.

In summary HighBrow can be considered to be informal, explicit and tacit, summarizing (without writing), hyperextensive and intensive, transient, and private. The highlights can easily be made public and group related for collaborative work.

3.1.6 HighBrow The Systems View

HighBrow is not a program but a system involving 25 classes and 3,292 lines of Java code. Each Java class serves a unique purpose in support of the highlighting and viewing process. This section describes HighBrow systems using the systems view of annotations developed by Cousins, Baldonado, and Paepcke (2000).

The BJEditorPane class is the *annotation writing platform* which is the annotator's primary interface. In HighBrow, the location of the annotations lies within the document itself. The size and capacity of the annotations can never exceed the size of the original document as they consist of highlighted text. The input medium is *digital ink* only without keystrokes.

The *annotation reading platform* is *collocated* with the writing platform and adds a separate pop-up summary (utilizing a JTextPane), and two split frame lists containing lists of previously highlighted material. Just as in the writing platform, the source document is located in the BJEditorPane, which contains keyword/phrase highlighting in bright yellow and context highlighting in a lighter pale yellow. The annotations are considered to be in the *overlaid* category because all of the annotations involve highlighting and the original document content is not altered. The pop-up context summary shows the keywords highlighted in bright yellow along with the context, which is not highlighted at all, since the summary, by definition, includes only context and keywords. Since the context summary is displayed in its own window, it is classified as *independent*. The context summary may be created from the annotations without retrieving the original document. The BJEditorPane provides *hideability* and is *searchable.* The summary does not provide *hideability* or *searchability*.

The object that is being highlighted is called the *annotation target.* In this dissertation the annotation target is the text within a Web document (i.e., it does not include images or sound) putting it in the *strict format dependence* area. While the highlighting appears to be *direct*, in actuality it is *indirect*. The highlights are not part of the document but rather an independent overlay of the original, thus the original document is never altered. The annotation target for HighBrow is considered to be *digital* rather than *physical*.

Annotations are the fourth component of the systems view, which includes *scope, structure, published, liveness,* and *stature*. The scope of HighBrow used in the experiments was considered to be *private* as sharing of the annotations was not allowed. Adding other levels of scope, specifically *workgroup* and *global,* is not a difficult process since the annotations stored in the database do carry the group information as well as personal information.

The highlighting action is *informal* since the reader is not required to fill out a form but the highlight itself is stored in a *formal* structure using a database table that includes the userid, the group id, date and time of creation, start and end location within the document and the actual text. The reason the location and the actual text is stored is so that future versions of HighBrow can detect changes in the document, notifying the user that the highlights are no longer current or correct.

With respect to *liveness*, the highlights in HighBrow can be considered *active*, since the annotations can be resized at any time or highlights can be changed from keyword to context to keyword with ease.

Highlights usually are defined with a *stature* of *fragment*, however with the introduction of context highlighting the stature may now be considered *document* since the ensuing summary may be much more readable and may contain a flow of thought rather than a series of isolated words.

HighBrow supports the fourth component of annotation systems, called *annotations correspondence* in the following ways: *mutability, anchoring granularity* and *technology*, and *robustness*,

Since many Web pages are becoming more dynamic, the content can be expected to change and is termed *mutable.* In its current state, HighBrow does not handle modifications to highlights when the underlying document changes. The highlights are based on location and if the page does change, the highlights may be highlighting data other than was intended. HighBrow could easily warn the user that the highlighting is no longer accurate, since the location within the annotation target and the content are maintained within the annotation itself. For purposes of the research to be described, the documents used were immutable.

The *anchoring granularity* of the highlights is at the *character* level. This gives HighBrow the advantage of being able to locate keywords or phrases within a document unlike other tools such as the browser favorites or bookmark lists, which are only able to locate information at the *document* level.

3.1.7 HighBrow Testing

The number of scenarios involving highlighting is large. Highlighting can be transferred from keyword to context and back to keyword easily, resized, deleted, etc.

The software underwent rigorous testing to make sure that all possibilities were addressed in order to prevent user confusion or frustration. Table 3.5 is a list of the conditions that were tested to ensure proper responses for all conditions. If any test failed, the code was modified and the entire set of tests was redone to ensure no new problems were introduced. As a result of this testing, no problems were encountered during the experiments, which, including the pilot tests, involved over 150 students.

3.2 HighBrow Versions

In order to control the experiments performed for this dissertation, multiple versions of HighBrow were produced with multiple sub-versions. Table 3.6 gives a summary of the versions, features, target group, and the number and purpose of any sub-versions. All participants used the HighBrow layout to eliminate any variation in the reading/annotating environment.

The different versions restricted navigation and highlighting capabilities to various extents. The sub-versions were used to help group class sections and also point to the content for the respective experiment. For Experiment One, both MidBrow and LowBrow pointed to the full document, for Experiment Three only MidBrow was used and it pointed to a series of Java Tutorials. NoBrow used in Experiment Two has three sub-versions each pointing to a separate document: the original, a context/keyword extract, and a keyword only extract.

Table 3.5 - Conditions Tested

Test	Pop-up Req'd	Action
New keyword	No	
Resize Keyword	Yes	
Add context over existing keyword	Yes	
Add keyword inside existing keyword	Yes	Previous keyword becomes context
Add keyword inside context	Yes	
Add keyword start of context	Yes	Extend context if necessary to include keyword
Add keyword end of context	Yes	Extend context if necessary to include keyword
Add second keyword inside	Yes	
Add second keyword start of context	Yes	Extend context if necessary to include keyword
Add second keyword end of context	Yes	Extend context if necessary to include keyword
Resize keyword inside of context	Yes	
Resize keyword start of context	Yes	Extend context if necessary to include keyword
Resize keyword end of context	Yes	Extend context if necessary to include keyword
Resize context with keywords inside	Yes	
Resize context start	Yes	
Resize context end	Yes	
Delete keyword inside	Yes	Causes context to become a keyword
Delete keyword start	Yes	Causes context to become a keyword
Delete Keyword end	Yes	Causes context to become a keyword
Delete Context	Yes	
Delete keyword inside context with multiple keywords	Yes	Context remains
Delete keyword start inside context with multiple keywords	Yes	Context remains
Delete Keyword end inside context with multiple keywords	Yes	Context remains

3.3 Summary

This chapter discussed the HighBrow system which is the primary interface used for the experiments discussed in detail in Chapter Four. The HighBrow architecture was compared to existing approaches and the HighBrow approach was discussed and justified. HCI aspects with respect to the layout and interface in general were explained.

The dimensions of annotation used in HighBrow were also presented as well as the systems view components that were utilized. The plan for testing the software was presented.

The following chapter describes the experiments that utilized the versions and sub-versions described in this chapter.

Table 3.6 - HighBrow Versions

Version Name	Features	Target Group	Number and Purpose of Sub-versions
HighBrow	Fully operational as described in this chapter	Used for development and testing, not used in the experiments	No sub-versions
MidBrow	Same as HighBrow except all references to URLs are removed	Used for experiment one, context group and experiment three all users	4 sub-versions to point to the appropriate homepage for each class/section and the longitudinal study
LowBrow	Similar to MidBrow except the context highlighting capability is removed	Used for experiment one, keyword group	3 sub-versions to point to the appropriate homepage for each class/section
NoBrow	Similar to MidBrow except all highlighting capability is removed	Used in experiment two for all groups	3 sub-versions to point to the appropriate homepage for the full document, context summary, and keyword summary

Chapter Four:

Experiments

Three separate experiments were conducted to determine if context highlighting is beneficial to annotators and readers, to examine context highlighting behavior including the number of words highlighted and time taken to highlight using context, and to see if context highlighting is useable and likable. The three experiments were as follows:

- Experiment One investigated the performance of annotators with respect to test scores, patterns of highlighting with respect to time and number of words highlighted, and whether or not the annotators found it useable and liked context highlighting.
- Experiment Two investigated the performance of readers of summaries, created from the work done by the participants in Experiment One, in preparation for a test. Experiment Two also tracks total preparation time, which is used to determine the efficiency of context highlighting.
- Experiment Three was a longitudinal study to determine highlighting habits and usage over time.

Each of these experiments will be discussed separately in detail below.

4.1 Experiment One

Obendorf (2003) explained that users would be unwilling to annotate without a clear benefit to themselves. HighBrow cannot construct context summaries if annotators are unwilling to annotate. Experiment One will determine if annotators benefit, using test performance as a metric, from context highlighting. Experiment One tracks the difference in keyword or phrase size (with respect to the number of words) as a result of context highlighting, and the amount of elapsed time between highlighting the keyword or phrase and the surrounding context. The survey at the end of Experiment One shows whether or not the context highlighting browser is useable and likeable.

4.1.1 Experiment One: Goals

The goals for Experiment One are as follows:

- To determine the effects on subsequent test performance for active readers who create context/keyword highlighting versus keyword only highlighting.
- To determine the number of words highlighted with context/keyword versus keyword only highlighting.
- To determine the usability and likeability of context/keyword highlighting.
- To provide data to be used in Experiment Two.

4.1.2 Experiment One: Questions

The following questions are answered by Experiment One:

- Does context/keyword highlighting influence test scores compared to keyword only highlighting?
- Are the number of words of the highlighted portion affected by context/keyword highlighting?
- Is the context/keyword highlighter usable and do people like it?

4.1.3 Experiment One: Method

In the first experiment, the participants were randomly assigned to one of two groups: context/keyword (also referred to as the context group) and keyword only. Both groups were required to read the same document. The participants were given a two-week period to read, study, and/or highlight the given document. The participants were given an assessment instrument consisting of a 20-item multiple-choice test. The test was the same for all participants. The participants were asked to complete a survey concerning the usability of HighBrow, the likeability of HighBrow features, and were allowed to offer free form comments about HighBrow.

4.1.4 Experiment One: Participants

The participants in the first experiment were enrolled in classes, consisting of two sections (early morning and early afternoon) of Introduction to Object Oriented Programming (OOP) and one section of File Structures, in the College of Computing, Engineering, and Construction at the University of North Florida, a medium size southeastern university. Participation was voluntary; however, each participant received extra credit for participating in and completing all phases of the study. Students not wishing to participate were given the opportunity to earn equivalent extra credit by writing a brief paper on the effects of stress using computers. No students elected this option.

To provide an incentive to perform well on the test, the participants were given extra credit based on their test scores, with students in the top third of each group receiving 10 points, middle third receiving 5 points, and the remainder receiving 3 points.

Participants from each class were randomly placed in the keyword group or the context group. Sixty-eight (68) students (33 context and 35 keyword) initially registered for the experiment. Of the 48 students (23 context and 25 keyword) who took the test, three participants were removed from the context group for failure to highlight keywords and context. One participant from the context group and two students from the keyword group were removed from the experiment because of inaccurate reporting due to data loss. The lost data were the result of the database quota being exceeded. Forty-two (42) students (19 context and 23 keyword) completed the experiment.

The groups were initially balanced by gender within the class/section, thus each group (context and keyword) had a near equal number of males and females, but within groups, the number of males and females differed. Table 4.1 shows the distribution of the participants who took the test. The highlighting and printing habits refers to how often the participant highlighted (no qualification was made regarding paper or digital) and printed (specifically Web) documents in the past. “Always” was omitted from the choices since it was felt that no one would do these activities all of the time, while there were cases of students who never highlighted or printed.

Thirty-nine (39) students completed the survey, 19 in the context/keyword group and 20 in the keyword only group. The loss of the three students in the keyword summary was due to attrition in the class.

Table 4.1 - Participant Demographics by Group

	Gender		Age Group				Highlighting Habit				Printing Habit			
Group	F	M	18-25	26-33	34-41	>41	never	rarely	often	most of the time	never	rarely	often	most of the time
Context	4	15	10	6	3	0	2	10	6	1	1	9	7	2
Keyword	4	19	16	3	2	2	3	9	9	2	5	11	6	1

4.1.5 Experiment One: Materials

The intervention for both groups in Experiment One required the participants to read excerpts from *The Effects of Computers on Workplace Stress, Job Security and Work Interest in Canada - December 2002* (see Appendix A) online (http://www.csee.usf.edu/~rjzucker/dissertation/cabdoc/EffectsofComputersonWorkplaceStressJobSecurityWorkInterest.htm). This document contained 76 paragraphs and 6073 words and focused on several issues relevant to the students participating in the study. These issues included the stress to learn new computer skills, to what extent computers affect work, job security related to computer work, the impact of computers on work interest, and the affects of workers with different attributes. This document was chosen because it was felt to be appropriate for use by participants in the first and second experiments. In the first experiment, the participants consisted primarily of Information Systems/Science students.

To ensure ecological validity in the first experiment, the students were allowed to study the document at the time and place of their choosing.

4.1.6 Experiment One: Instruments

For this experiment, the following instruments were used: an online registration form, downloadable installation software, modified versions/sub-versions of HighBrow, a test, and an online survey.

4.1.6.1 Registration

In order for a user to use the HighBrow system, the user is required to register using an anonymous userid. The userid is necessary to ensure that highlights are associated with the correct users and to keep track of timing by users. An important design consideration is to allow registration anywhere (at home, work, school, etc.), therefore a Web based registration form was created using Java Server Pages (jsp). There are three separate registration pages, one for each course and/or section, in order to keep track of participants by course/section. From the user standpoint, all registration screens are identical in appearance, the internal difference is in the group id (IOam, IOpm, and FS) assigned to the registrant. The registration screen allows a participant to create a userid and enter demographic information: gender, age group, as well as highlighting and printing habits. Appendix B shows the registration screen used for Experiment One.

For Experiment One, the participants were allowed to read and highlight in an unmonitored environment; therefore, they were required to download the browser onto the machine[s] of their choice.

Since the experiment consisted of two groups, context/keyword highlighters and keyword only highlighters, participants were required to highlight using one of the two versions of the browser. This meant that upon successful completion of the registration process, the registration screen would point the registrants to the proper site to download the correct version and sub-versions of HighBrow. To ensure that the groups were balanced, the registration process used a random number generator and gender within the class/section to determine which group to assign to the participant. Odd numbered registrants in a given section, within gender, were randomly assigned to one group and even numbered registrants were assigned to the alternate group that would allow balance.

4.1.6.2 HighBrow Installation

When the registrant successfully registers, a link to the installation document is presented. When the link is opened, the registrant is presented with a page containing detailed instructions on the installation process (see Appendix C). Since there are two different versions of HighBrow, there are two installation processes differing only by the link to the zip file containing the software to be used in the experiment.

Each zip file contains: the version of HighBrow required for that group; two batch files, one to build the environment and one to run the program; a program to detect the version of Java used and either set the proper directory or notify the user to download

Java 1.5. The zip file contained the Oracle database Application Program Interface (ojdbc14_g.jar) to provide database connectivity.

4.1.6.3 HighBrow

For this experiment, two sub-versions of HighBrow were involved: one to allow context/keyword highlighting, called MidBrow, and the other, called LowBrow, to allow keyword only highlighting. Both MidBrow and LowBrow were created to allow link following only, thus controlling the browsing experience, and do not display the URL address in order to reduce the opportunity for students to study the material with their usual browser. All of the other functionality of HighBrow was maintained in MidBrow and LowBrow.

4.1.6.4 Test

The test contained 20 multiple-choice items (see Appendix D) with the number of choices ranging from two to five. The principal investigator and another instructor each were responsible for developing a test. The final test was constructed by consolidating questions from the two tests which were developed independently. The scores from the test had a possible range from a minimum of zero to a maximum of 20. The types of questions varied from simple knowledge based questions, for example:

The “Effects of Computers” study was performed in

a) Canada
b) Mexico
c) U.S.A.

to questions requiring comprehension, for example:

> The association to which work is affected by computers and the extent/frequency of computer usage is:
> *a) Directly proportional.*
> b) Inversely proportional.
> c) No pattern exists in relation.

The test was given in the last twenty minutes of each class at the end of the two-week study period so that the participants in each section were given the test at the same time.

4.1.6.5 Usability Survey

An online usability survey (Appendix J) contained 21 items. The first nine items consisted of Likert five point scale usability questions (with values of 5-strongly agree, 4-agree, 3-neutral, 2-disagree, 1-strongly disagree), and 0 if not applicable. There were 11 items dealing with perceptions of capabilities using a Likert three point scale (2-like, 1-dislike, 0-neutral/no opinion). The last item was an optional open-ended future enhancements comment box. The survey was evaluated by three independent reviewers for content and wording.

4.1.7 Experiment One: Procedures

The author performed all data gathering. Potential participants were given an explanation of the responsibilities and risks regarding participation and were required to sign a Human Research Informed Consent Form (see Appendix E). The participants were also given guidelines, both orally and in written form, describing what was expected to be learned from the reading. Students in the context/keyword group were told that if they chose to highlight, that they must also highlight context. Students in both groups were told that highlighting was not required. Both groups were present when the instructions were given.

Both groups were told orally, via e-mail, and via Blackboard the registration site address and how to register. The registration site randomly assigns either the context/keyword or keyword only group based on gender and class section, thus each group (context and keyword) had a near equal number of males and females, but within classes, the number of males and females differed. Upon successful registration, the student is given a link to the proper installation procedures for their respective group.

After the correct version of HighBrow is installed and executed, a homepage for that specific group is displayed. The homepage contains a thank you message for participating and briefly describes the type of questions that would be encountered in the test. The homepage also included two links, one pointing to instructions on how to use the highlighting browser and one pointing to the document that is the basis for the test. There are two unique instruction pages, one for context highlighting/keyword highlighting and one for keyword only highlighting (see Appendix F and Appendix G). The instruction page is the only guidance that the participants receive regarding the usage of the highlighting browser. The students were allowed two weeks to use the features of HighBrow including printing the original document, printing the summary, copying the summary to another document, and/or highlighting as they saw fit.

HighBrow does record a history of mouse clicks, which are called events, with the userid, date and time, function, and when appropriate the start and end location of the highlight. Since this study period was not in a controlled environment, the timing reported by HighBrow cannot be assured to be an accurate measure of time on task (e.g., a student could open the site and be interrupted by a phone call, dinner, etc).

Students in the context group, if they chose to highlight, were required, as a condition of participation, to highlight both keywords and context. To ensure compliance with the requirement, a special Web site was set up to provide confirmation that each keyword contained some context. Figure 4.1 provides a sample of the output from the verifier when context or a keyword is missing in a document. The URL document is intentionally truncated to "putersonWorkplaceStressJobSecurityWorkInterest" to prevent participants from accessing the Web site without HighBrow. If the highlighting was done correctly the site would report: "Number of missing entries: 0". This site merely verified that the context/keyword requirement was adhered to; it did not impose any requirement that highlighting must be present. Despite the availability of this tool, three participants were deemed ineligible to participate for failing to highlight context surrounding a keyword.

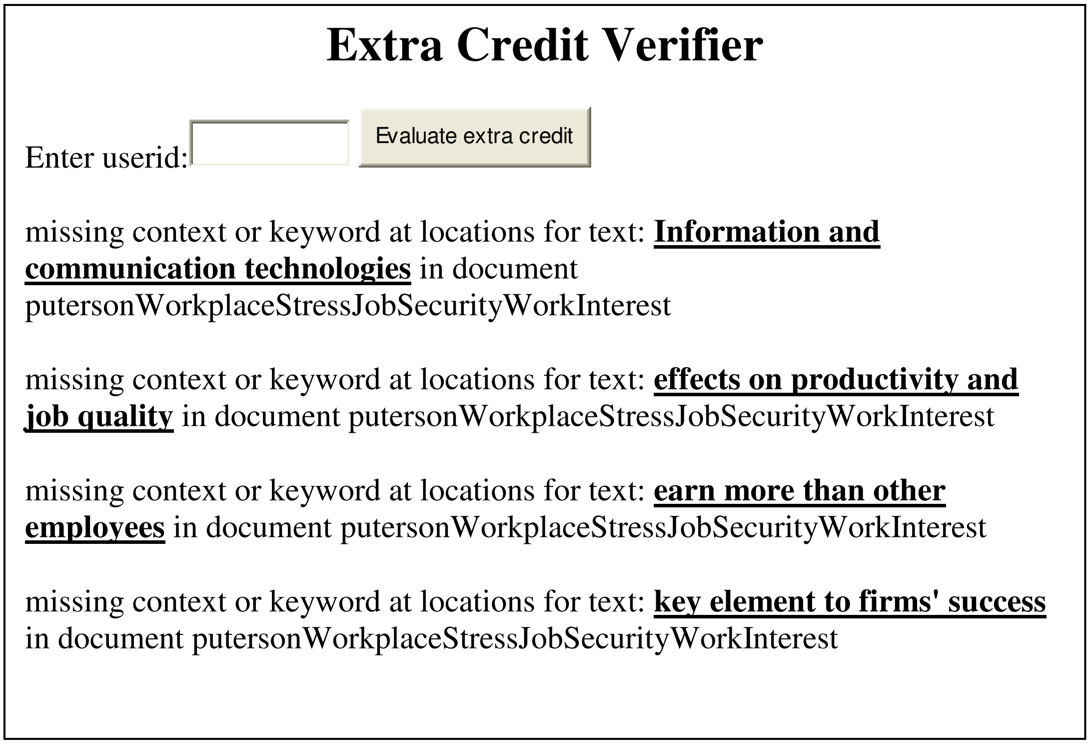


Figure 4.1 - Screenshot of Context Verifier

At the end of the study period, the test was administered in each of the three classes. Approximately 20 minutes were allotted to take the test. Time was not considered a limiting factor in test performance as no student took more than 15 minutes to complete the test. The scores were tabulated and entered into a spreadsheet for evaluation (see Appendix H).

The students were then asked to complete the usability survey online. Thirty-nine students completed the survey.

4.1.8 Experiment One: Results

In order to answer the first question, "*Does context and keyword highlighting influence test scores compared to keyword only highlighting?",* the test scores were evaluated to determine difference between the context/keyword group and the keyword only group. An independent-samples *t* test was conducted to evaluate the hypothesis that scores of participants highlighting context/keyword differed from scores of participants highlighting keywords only. The test scores for the context/keyword group on the average (*M*=12.42, *SD*= 3.01) appeared to be higher than the average test scores (*M*=11.35, *SD*=2.44) for the keyword only group (see Table 4.2). The boxplot in Figure 4.2 shows the distribution of scores for the two groups. with homogeneity of variances of differences assumed (see Table 4.3), the test was not significant (see Table 4.4) for the double tailed test ($t(40)=1.28$, $p=0.22$). The 95% confidence level ranging from -0.625 to 2.771 is quite wide. Cohen's effect size $d=0.40$ indicates a small to medium effect which suggests significance may exist given a sample size greater than 64 for each group.

Table 4.2 - Mean and Standard Deviation of Test Scores by Group

	Group	N	Mean	Std. Deviation	Std. Error Mean
Score	Context	19	12.42	3.006	.690
	Keyword	23	11.35	2.442	.509

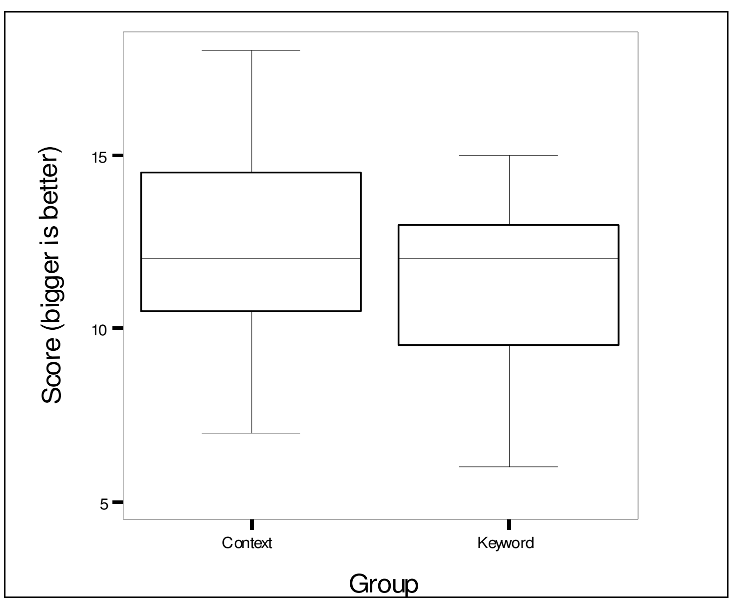


Figure 4.2 - Test Score Boxplot

Table 4.3 - Experiment One Homogeneity of Variances of Differences of Test Scores

Levene's Test for Equality of Variances	
F	Sig.
.467	.498

Table 4.4 - Test Score, Independent Samples Test

		t test for Equality of Means						
							95% Confidence Interval of the Difference	
		t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Score	Equal variances assumed	1.277	40	.22	1.073	.840	-.625	2.771

The second question, "*Are the number of words of the highlighted portion affected by context highlighting?*" relates to the number of words highlighted. The number of words highlighted is important for two reasons: fewer words lead to improved recall (Lorch, Pugzles-Lorch, & Klusewitz, 1995), and fewer words make indices easier to read. In Lorch's experiment, underlining was used to annotate the document; however Lorch defined underlining as typographical cues or signals which include "capitalization, italics, boldface, and color variation".

An independent samples *t* test was conducted, using the number of keywords highlighted as the dependent variable and the highlighting condition, (context/keyword and keyword only) as the independent variable. The mean number of keywords highlighted by the keyword only group (*M*=853.17, *SD*= 897.65) was significantly higher than the mean number of keywords highlighted by the context/keyword group (*M*=303.95, *SD*=428.78) as shown in Table 4.5. The boxplot in Figure 4.3 shows the distribution of keywords for the two groups. Homogeneity of variances of differences was not assumed (see Table 4.6). The 2-tailed test is significant, $t(32.78)=2.60$, $p=0.01$ (see Table 4.7).

Table 4.5 - Mean and Standard Deviation of Keywords by Group

	Type	N	Mean	Std. Deviation	Std. Error Mean
Keywords	Keyword	23	853.17	897.645	187.172
	Context	19	303.95	428.779	98.369

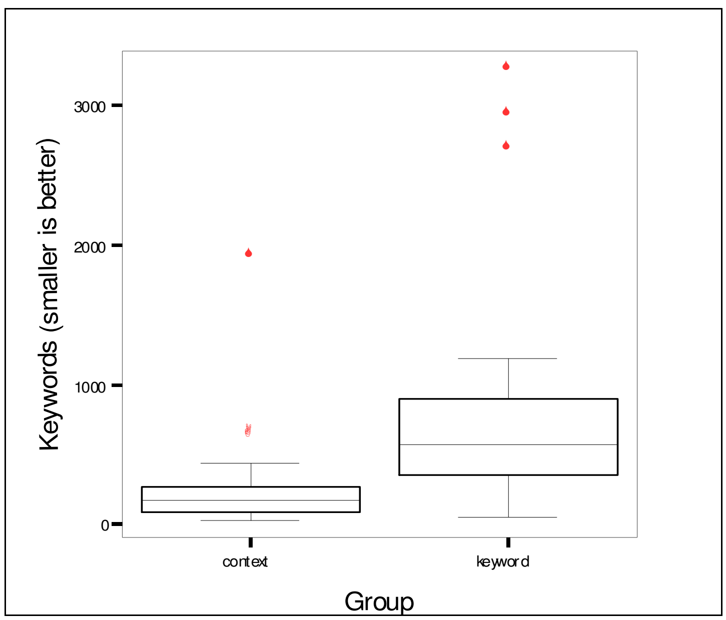


Figure 4.3 - Number of Keywords by Group Boxplot

Table 4.6 - Homogeneity of Variances of Differences of Keywords

Levene's Test for Equality of Variances	
F	Sig.
4.317	.044

Table 4.7 - Number of Keywords, Independent Samples Test

		t test for Equality of Means						
		t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
							Lower	Upper
Key words	Equal variances not assumed	2.597	32.775	.014	549.227	211.447	118.923	979.530

If more words aid readability, then it would be beneficial to highlight more words for readability of the context summary. In order to compare the number of words highlighted with respect to context, an independent samples *t* test was conducted, using the number of keywords highlighted in the keyword only group and the number of context words highlighted in the context/keyword group as the dependent values and the keyword highlighting group and the context/keyword group as the independent variable. The number of words highlighted by the context/keyword group (*M*=1614.58, *SD*=1126.58) on the average were more than the number of words highlighted by the keyword only group (*M*=853.17, *SD*= 897.65) as summarized in Table 4.8. Figure 4.4 shows the distribution of keywords for the two groups. With homogeneity of variances of differences assumed (see Table 4.9), the result from the independent samples *t* test was significant ($t(40)=-2.44$, $p=0.02$). The independent samples *t* test result is shown in Table 4.10. The 95% confidence level ranging from -1392.435 to -130.376 is quite wide.

Table 4.8 - Mean and Standard Deviation of Words Highlighted by Type (Context/Keyword Group)

Type	N	Mean	Std. Deviation	Std. Error Mean
Keyword	23	853.17	897.645	187.172
Context	19	1614.58	1126.576	258.454

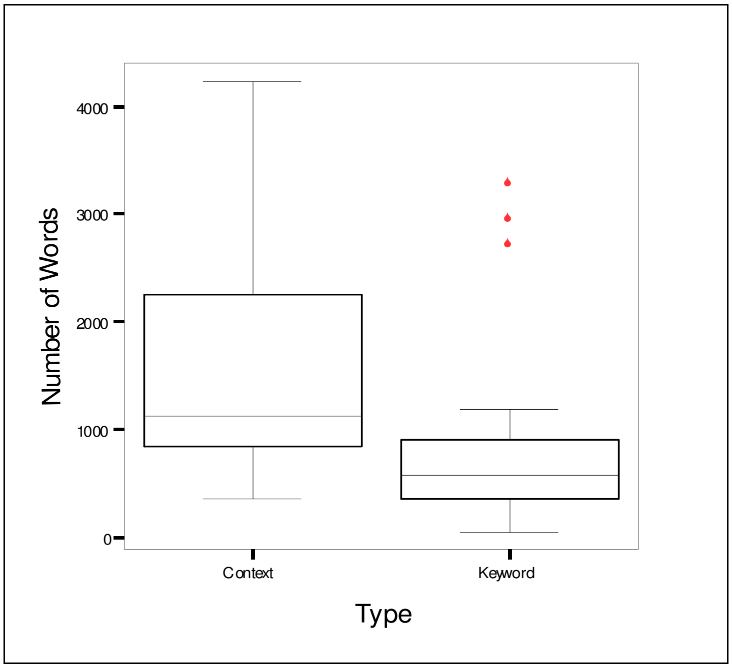


Figure 4.4 - Words Highlighted within Context/Keyword Group, Boxplot

Table 4.9 - Homogeneity of Variances of Differences of Number of Highlighted Words

Levene's Test for Equality of Variances	
F	Sig.
2.108	.154

Table 4.10 - Number of Context Words, Independent Samples *t* Test

		t test for Equality of Means						
							95% Confidence Interval of the Difference	
		t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Context	Equal variances assumed	-2.439	40	.019	-761.405	312.225	-1392.435	-130.376

To see how long the additional highlighting (context added to keyword or keyword added to context) took, the timestamps between the keyword and the context highlights were selected from the final highlighting data. Initially the historical data were used, however, the historical data contained events that would confuse the data, for example adding and deleting, or modifying existing keywords/context, whereas the final highlighting data showed the final result of the highlighting action. It is important to note that the timings reported here do not necessarily indicate the time on task, only the time between the time the context was highlighted and the time the keyword was highlighted. In some cases, days passed between the time a keyword or context and its corresponding context or keyword were highlighted. Table 4.11 shows the average time in seconds by user and the number of keyword highlighting events (not the number of keywords) by user. The timing results for all keyword highlighting events are highly skewed to the right with the mean being 77,927 seconds (21 hours, 38 minutes, and 47 seconds) and the median being 10 seconds.

Table 4.11 - Mean Time Between Keyword and Context Highlighting Event (All Events)

User	Mean (seconds)	Number of Events	Std. Deviation
	73409.69	35	302512.599
	9.69	123	10.649
	7493.39	36	44812.437
	94.62	13	172.243
	13.06	135	18.446
	10.94	33	11.937
	7.66	29	4.073
	52183.00	46	238979.933
	8.68	118	12.280
	22092.48	21	101184.537
	13.61	49	14.922
	11.39	62	18.777
	7.78	65	6.902
	230829.49	423	294584.355
	102710.47	43	379404.261
	18.52	52	22.017
	80.81	31	390.349
	15.50	30	22.031
	14.23	39	16.880
Total	77927.00	1383	214427.538

Table 4.12 presents the timing results assuming that the time difference in highlighting keywords and highlighting the corresponding context (or vice versa) would take less than 180 seconds, eliminating any events taking more than 180 seconds. The assumption is based on the thought that if more than 3 minutes had passed, then the user interrupted the highlighting process. The events less than or equal to 180 seconds were also skewed to the right, with the mean being 19.1 seconds and the median being 9 seconds. Figure 4.5 shows that one user averaged more than 30 seconds while most users averaged less than 15 seconds between highlighting keywords and context.

Table 4.12 - Mean Time Between Keyword and Context Highlighting Event (Event Delta <=180 Seconds)

User	Mean (seconds)	Number of Events	Std. Deviation
	7.85	33	2.181
	9.69	123	10.649
	9.48	33	19.043
	24.82	11	21.372
	13.06	135	18.446
	10.94	33	11.937
	7.66	29	4.073
	7.84	43	7.546
	8.68	118	12.280
	12.20	20	7.675
	13.61	49	14.922
	11.39	62	18.777
	7.78	65	6.902
	49.75	240	45.505
	15.13	39	18.003
	18.52	52	22.017
	10.73	30	12.646
	15.50	30	22.031
	14.23	39	16.880
Total	19.17	1184	28.932

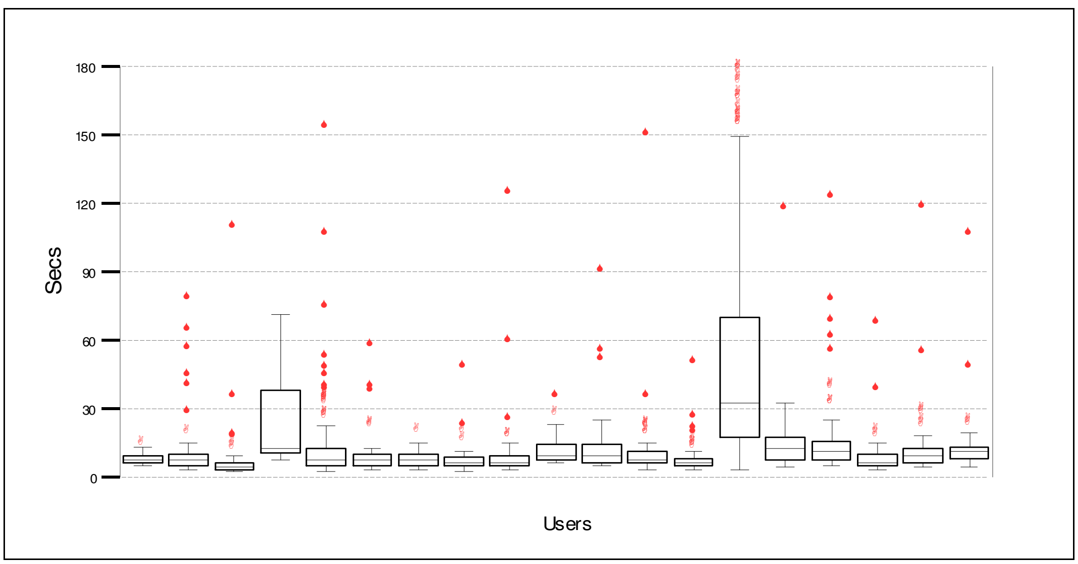


Figure 4.5 - Boxplot of Keyword and Context Highlighting Event (Event Delta <=180 Seconds)

The last question, "*Is the context highlighter usable and do people like it?*" is answered using the responses from the participants in the context/keyword group to the on-line survey. Table 4.13 presents the survey results.

Table 4.13 - Context/Keyword Group Usability Responses

	N	Minimum	Maximum	Mean	Std. Deviation
Easy to install	19	4	5	4.74	.452
Easy to use	19	4	5	4.58	.507
Loading web pages was fast	19	0	5	3.89	1.197
Highlighting was fast	19	4	5	4.68	.478
Easy to learn	19	3	5	4.37	.684
Context was beneficial	19	3	5	4.47	.697
I would use HighBrow	19	2	5	4.16	.765
Liked the Layout of Components	19	3	5	4.11	.658
Positive Overall Experience	19	3	5	4.47	.612

Legend

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	n/a
Response	5	4	3	2	1	0

The top five context/keyword group responses indicated that with HighBrow: highlighting was fast, easy to install, easy to use, context was beneficial, and the overall experience was positive. All of these questions received a mean rating ranging from 4.47 and 4.74 which is between agree and strongly agree. The lowest rating was given to the loading pages was fast question, with a score of 3.89 which is between agree and neutral. Responses to the remaining questions were above 4.0, which is between agree and strongly agree.

Table 4.14 summarizes the survey results for context/keyword participants with respect to the capabilities of Highbrow. The capabilities the context/keyword group liked most were: ability to highlight, ability to view a summary, ability to locate highlights this page, and ability to delete highlights. There was 100% agreement that the ability to highlight was liked by this group.

Table 4.14 - Context/Keyword Group Likeability Responses

	N	Minimum	Maximum	Mean	Std. Deviation
Ability to Highlight	19	2	2	2.00	.000
Ability to Modify existing Highlights	19	0	2	1.63	.761
Ability to Delete Highlights	19	0	2	1.74	.653
Ability to Hide/Show Highlights	19	0	2	1.37	.955
Ability to Locate keywords this Page	19	0	2	1.79	.631
Ability to Locate keywords other pages	19	0	2	1.58	.838
Ability to View Summary	19	0	2	1.89	.459
Ability to Print Summary	19	0	2	1.47	.905
Ability to Copy Summary	19	0	2	1.47	.905
Ability to Print Document with Highlights	19	0	2	1.37	.955
Ability to Delete all Highlights this page	19	0	2	1.63	.761

Legend

	Like	Dislike	No Opinion
Response	2	1	0

The keyword only group also took the survey and their responses are shown in Table 4.15. The top four keyword only group mean ratings indicated that with HighBrow: highlighting was fast, easy to install, easy to use, and the overall experience was positive ranged from 4.10 to 4.40 which is between agree and strongly agree. The lowest rating, M=1.65 (between disagree and strongly disagree), went to context was beneficial. The keyword only group did not get a context/keyword summary but a summary of the highlights, which contained the label *Context Summary*. In general, the responses from the keyword only group to the usability questions show slightly less agreement than the responses from the context/keyword group.

Table 4.15 - Keyword Only Group Usability Responses

	N	Minimum	Maximum	Mean	Std. Deviation
Easy to install	20	3	5	4.40	.681
Easy to use	20	2	5	4.15	.813
Loading web pages was fast	20	0	5	3.70	1.302
Highlighting was fast	20	4	5	4.45	.510
Easy to learn	20	0	5	3.90	1.252
Context was beneficial	20	0	5	1.65	1.927
I would use HighBrow	20	2	5	3.85	.875
Liked the Layout of Components	20	2	5	3.55	.945
Positive Overall Experience	20	2	5	4.10	.641

Legend

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	n/a
Response	5	4	3	2	1	0

Table 4.16 shows the survey results with respect to the liked capabilities of features by the participants in the keyword only group. The capabilities the keyword only group liked most were: ability to highlight, ability to view summary, ability to locate keywords this page, and ability to hide/show highlights.

Table 4.16 - Keyword Only Group Likeability Responses

	N	Minimum	Maximum	Mean	Std. Deviation
Ability to Highlight	20	1	2	1.95	.224
Ability to Modify existing Highlights	20	0	2	1.65	.587
Ability to Delete Highlights	20	0	2	1.60	.598
Ability to Hide/Show Highlights	20	0	2	1.80	.616
Ability to Locate keywords this Page	20	0	2	1.80	.616
Ability to Locate keywords other pages	20	0	2	1.30	.979
Ability to View Summary	20	0	2	1.80	.616
Ability to Print Summary	20	0	2	1.30	.979
Ability to Copy Summary	20	0	2	1.10	1.021
Ability to Print Document with Highlights	20	0	2	1.50	.889
Ability to Delete all Highlights this page	20	0	2	1.35	.813

Legend

	Like	Dislike	No Opinion
Response	2	1	0

In summary, there was a very positive agreement in the responses of both groups to the speed of highlighting, and ease of installation and use. The responses from both groups indicated both groups liked the ability to highlight, to locate keywords this page and to view the summary. The least positive from both groups (*M*=1.65) was in response to the context was beneficial question. This response was from the keyword only group and was to be expected.

4.1.9 Experiment One: Summary

Experiment One showed that the mean scores for this experiment were higher between the context/keyword annotating group and the scores for the keyword only annotating groups, however the difference was not statistically significant. People who highlight with context, with respect to the mean, highlight significantly fewer keywords (M=303.95) than people who highlight with keywords alone (M=853.17). However people who highlight context, with respect to the mean, highlight more total words (M=1614.58) than the people who highlight keywords only (M=853.17). Both versions of Highbrow, MidBrow and LowBrow, were found to be usable and well liked by the participants. The biggest difference in usability was the responses to the "Context was beneficial" question, the context/keyword group mean rating was 4.47 (between agree and strongly agree) and the keyword only group mean rating was 1.65 (between disagree and strongly disagree). The keyword only group did not have a true context summary but rather a summary of highlighted text which was called a context summary.

4.2 Preparation for Experiment Two, the Context and Keyword Extraction Process

In order to perform Experiment Two it was necessary to extract a set of keywords which would be representative (in terms of number and specific words) of highlighting done by participants in the keyword condition of Experiment One. It was also necessary to extract a set of context and keywords that would be representative (in terms of number, specific words, and appearance (i.e., eliminating keywords with no associated context or context without associated keywords)) of highlighting done by participants. The resulting summaries were used in the second experiment for groups two and three (context and keyword highlights, and keyword only highlights).

To facilitate the extraction process, a Java program was written to extract context and/or keywords and provide an objective extraction of content. The following is a description of the algorithms used by the extraction programs to extract keywords only and context and keywords.

4.2.1 Keyword Only Extraction

The keyword only extraction program is simpler than the context and keyword extraction program. The keyword extraction program processes each keyword highlighted by each participant, accumulating the number of keywords highlighted in the given document, and accumulating the number of agreements for each word (among participants). The granularity of text chosen was a single word, so if a partial word is mistakenly highlighted, the entire word is used (i.e., no partial words are used).

As a first step, the summary should contain approximately the same total number of highlighted words as the mean number of highlighted words from Experiment One. The program calculates the mean number of words highlighted by counting all of the highlighted words and dividing by the number of participants. In this experiment, 23 participants highlighted a total of 19,869 keywords, resulting in an average of 856 highlights per participant. This mean represents a target number for the number of words to include in the keyword summary.

Next, the specific words to be included must be chosen. This may be done by identifying those words which were highlighted most frequently by Experiment One participants. The program counts the number of words agreed upon by group as graphed in Figure 4.6. Starting with the group containing the largest agreement and working down to the group containing the smallest agreement (usually but not always one), the program subtracts the number of words per group from the target, thus obtaining the words with the highest agreement first, the next to the highest next, until all of the words from the target will fit into a group.

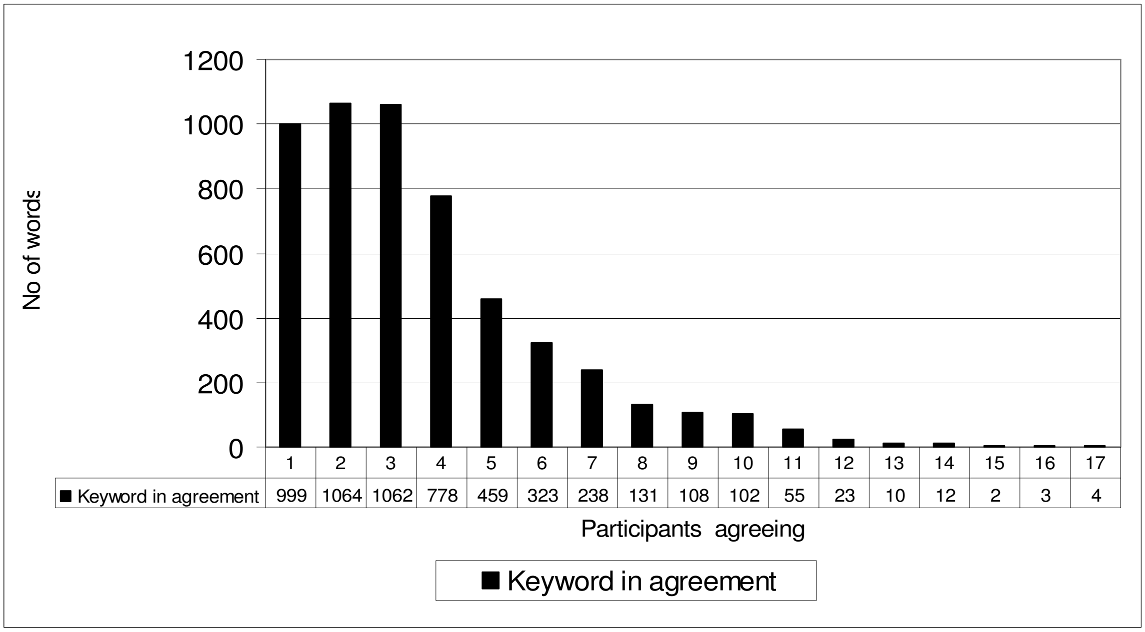


Figure 4.6 - Participant Agreement Groups for the Keyword Only Group

The process continues until the point where the inclusion of words would result in a set which exceeded the target size. At this point, a decision must be made as to how to allocate the words from that group. One way is to randomly select words to be used until all of the words have been allocated, another way is to semantically process the highlighted words (e.g., removing indefinite articles (a, an, the) and conjunctions (and, or, nor, etc)). In these experiments, the extraction program determines if the number of words in the final group is more than half of the remaining words. If so, then all of the words in that group are extracted (see Table 4.17). If the words remaining are less than half of the group size, then none of those words are extracted. This method was chosen so that words were not randomly and arbitrarily deleted from a given agreement group possibly destroying context within agreement groups.

Table 4.17 - Agreement Groups and Extracted Words for the Keyword Extraction Process

Participants in agreement	Number of words in group	Number taken from group	Number remaining (from 856)	Comments
17	4	4	852	All words taken
16	3	3	849	All words taken
15	2	2	847	All words taken
14	12	12	835	All words taken
13	10	10	825	All words taken
12	23	23	802	All words taken
11	55	55	747	All words taken
10	102	102	645	All words taken
9	108	108	537	All words taken
8	131	131	406	All words taken
7	238	238	168	All words taken
6	323	323	0	Since 168 > half of 323 all were taken
5	459	459	0	None taken
4	778	778	0	None taken
3	1062	1062	0	None taken
2	1064	1064	0	None taken
1	999	999	0	None taken

When the extraction method was employed, the number of keywords in the final extraction was 1,011, while the target was 856. This overage of 155 words was the result of 168 words remaining to be extracted which was more than half of the 323 words in the last agreement group. The number of participants in agreement was six or more people.

4.2.2 Content and Keyword Extraction

The context and keyword extraction is similar to the keyword only extraction. The difference in the two extraction methods lies in the fact that the context group double highlighted some words (once for keyword and once for context). Another important aspect of the extraction process is to ensure that the extracted document must have the same appearance as an average context and keyword document, that is, each context must contain a keyword and each keyword must have an associated context.

The first step is to determine the target number of words to include in the extract. Table 4.18 shows how the target was obtained for this experiment. Nineteen participants highlighted 36,561 words, both context words and keywords. In order for the final document to contain the correct number of words, the 5,775 keywords were removed from the total since the keywords have already been included in the context resulting in 30,786 words. Dividing this number by the 19 participants results in a target of 1,620 words.

Table 4.18 - Determining the Target Words for Extraction

Total words highlighted	36,561
Keywords	5,775
Total words highlighted-keywords	30,786
Target (average from 19 participants)	1,620.32

Once the target is obtained, the words are extracted in a manner similar to the keyword only group. Table 4.19 shows how the extraction process selected the initial group of words. In this case, the group of participants in agreement of eight had 373 words with only 222 remaining from the average. Since 222 words are more than half of the group's 373, all of the words were used, resulting in an overage of 151 with a final count of 1,771 words.

Table 4.19 - Agreement Groups and Extracted Words for the Initial Context Extraction Process

Participants in agreement	Number of words in group	Number taken from group	Number remaining (from 1620)	Comments
27	2	2	1618	All taken
26	1	1	1617	All taken
25	1	1	1616	All taken
24	1	1	1615	All taken
23	4	4	1611	All taken
22	3	3	1608	All taken
21	15	15	1593	All taken
20	11	11	1582	All taken
19	17	17	1565	All taken
18	28	28	1537	All taken
17	49	49	1488	All taken
16	53	53	1435	All taken
15	62	62	1373	All taken
14	84	84	1289	All taken
13	100	100	1189	All taken
12	182	182	1007	All taken
11	174	174	833	All taken
10	266	266	567	All taken
9	345	345	222	All taken
8	373	373	0	Since 222 is greater than half of 373 all were taken
7	602	602	0	none taken
6	462	462	0	none taken
5	600	600	0	none taken
4	740	740	0	none taken
3	828	828	0	none taken
2	627	627	0	none taken
1	294	294	0	none taken

To ensure that the extracted document has the appearance of the average context and keyword document, the extract program must examine the highlighted phrases, eliminating phases that are keywords with no associated context or context without associated keywords. Keywords are identified and as long as there is agreement that the keyword is important by at least half of the participants, the keyword is included.

Because of this keyword identification process, it is possible for the entire keyword/phrase to be the context, or for there to be context without at least 50% agreement. The result would be that some context would be keyword[s] only or context only. To prevent this from happening, the context is then validated by searching for at least one keyword in the context area and at least one context word in the keyword area.

As a result of the validation process, 232 words were eliminated, causing the context extraction to contain 1,539 words, 81 words less than the original target of 1,620. Attempts were made to include additional groups to make up for the loss, but as the groups have less agreement, the number of words in the group grow (see Figure 4.7). In this experiment, including the agreement group with seven participants agreeing would add 602 words (see Table 4.19), forcing the number of words highlighted to go over the target.

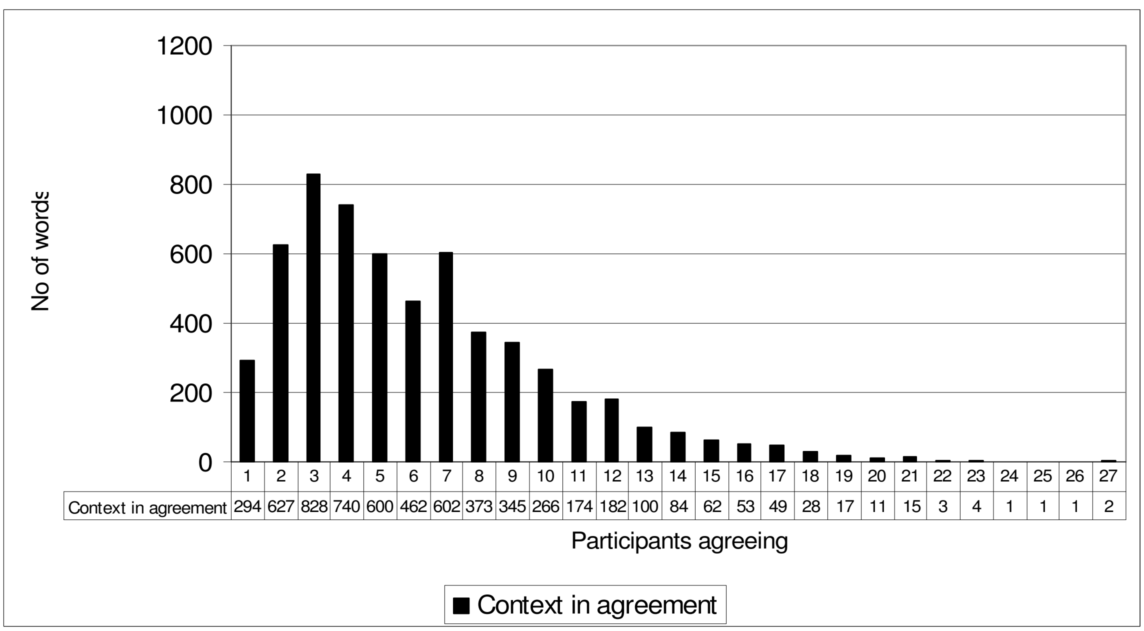


Figure 4.7 - Context Agreement Groups

4.3 Experiment Two

Experiment Two will determine if passive readers benefit from context highlighting. For this experiment, passive readers are defined to be readers without the ability to highlight, however they were not true passive readers since they were allowed to take notes during their reading.

4.3.1 Experiment Two: Goals

The goal for Experiment Two is as follows:

- To determine the benefits to passive readers who read a context highlighting summary, as compared to readers who read the full document or a keyword highlighting summary using scores, time, and efficiency as dependent variables.

4.3.2 Experiment Two: Questions

The following questions are answered by Experiment Two for passive readers:

- Does reading a complete document, a document with context/keywords, or a document with keywords only improve test scores?
- Does reading a complete document, a document with context/keywords, or a document with keywords only reduce study time?
- Is study time and test performance together enhanced by reading a complete document, a document with context/keywords, or a document with keywords only?

4.3.3 Experiment Two: Method

In the second experiment, the participants were randomly assigned to one of three groups and were required to read the respective version (complete document, context/keyword summary, or keyword summary only) of the same document. The participants were given a two-week period to read/study the given document. The participants were given an assessment instrument consisting of a 20-item, multiple-choice test. The test was the same for all participants and was the same as the test for the Experiment One students.

4.3.4 Experiment Two: Participants

The participants in the second experiment were enrolled in classes in the College of Arts and Sciences at the University of North Florida consisting of three sections (early and late morning and late afternoon) of Social Psychology and two sections of Cognitive Psychology (early morning and late afternoon). Participation was voluntary, and each participant received extra credit for participating in and completing both phases of the study. The total number of persons agreeing to participate was 60.

To provide an incentive to perform well on the test the participants were given monetary awards of $20, $15 and $10 for the top three scores for each of the three groups.

Participants from each class were randomly placed in one of the three groups by the first person selecting a card from a group of three cards, the second person selecting from a group of the two remaining cards, and the third simply being placed in the remaining group. The process was then repeated for each additional group of three participants, thus each group consisted of 20 participants. Despite the fact that no attempt was made to balance the groups by gender, Table 4.20 shows that the participants were evenly distributed. As in Experiment One, the highlighting and printing habits refers to how often the participant highlighted (no qualification was made regarding paper or digital) and printed (specifically Web) documents in the past. “Always” was omitted from the choices since it was felt that no one would do these activities all of the time, while there were cases of students who never highlighted or printed.

Table 4.20 - Experiment Two Participant Demographics by Group

	Gender		Age Group				Highlighting Habit				Printing Habit			
Group	F	M	18 - 25	26 - 33	34 - 41	> 41	never	rarely	often	most of the time	never	rarely	often	most of the time
Document	16	4	14	5	1	0	2	8	8	2	1	8	7	4
Context Summary	15	5	14	3	1	2	1	12	5	2	4	8	6	2
Keyword Summary	13	7	16	3	0	1	3	5	9	3	3	8	7	2

Sixty students successfully completed the second experiment, with 20 students in each group.

4.3.5 Experiment Two: Materials

The intervention for the groups in Experiment Two required the participants to read three varying degrees of excerpts from the same document used in Experiment One, *The Effects of Computers on Workplace Stress, Job Security and Work Interest in Canada - December 2002* online (http://www.csee.usf.edu/~rjzucker/dissertation/cabdoc/EffectsofComputersonWorkplaceStressJobSecurityWorkInterest.htm).

The first condition was the entire document. The second condition's document was created by using extraction software (see section 4.2.2) using the highlights created by the students who participated in Experiment One's context/keyword group. The context/keyword summary document contains 1,539 words. The document is arranged with a line break between groups of highlighted context. The highlighted keywords are shown in bright yellow. For the final condition, the keyword only excerpt was also created by using extraction software (see section 4.2.1) using the highlights created by the keyword only group in Experiment One. The keyword only extract contains 1,011 words. Since this extract contains only keywords, they are not highlighted. Similar to the context/keyword summary, the keyword only summary contains line breaks after each group of highlights.

To ensure temporal validity, the students were required to study the document assigned to their group (full document, context/keyword summary document, and keyword only summary document) in a controlled laboratory environment. The students were permitted to take breaks or to spread out the study time over multiple sessions. Four students did break up their study time into two periods.

4.3.6 Experiment Two: Instruments

For this experiment, the following instruments were used: an extraction program, installation, registration, the NoBrow version of HighBrow, and a test.

4.3.6.1 Registration

This screen was similar to the registration screen used in Experiment One except that the student was provided a dropdown menu to enter that student's group number.

4.3.6.2 Installation Procedure

Since the experiment was conducted in a controlled laboratory environment, the three sub-versions of HighBrow were preloaded onto the machines by the investigator.

4.3.6.3 HighBrow

To preserve continuity in the environment for both Experiment One and Experiment Two, sub-versions of HighBrow, each called NoBrow, were developed containing the same look and feel for reading a document as Experiment One. It was also beneficial to use NoBrow for data gathering since NoBrow could accurately record the study time. The three versions of NoBrow pointed to three separate homepages: one with a link to the full document, a second with a link to point to the context/keyword summary document, and the last with a link that pointed to a keyword only summary. Other than the links, the homepages (see Appendix I) were alike in every way.

None of the NoBrow versions support highlighting in any form, context or keyword, since this was an experiment restricted, with respect to highlighting, to passive readers.

As in Experiment One, NoBrow is restricted to allow link following only (thus controlling the browsing experience) and is not able to display the URL address in order to reduce the opportunity for students to study the material with their usual browser.

4.3.6.4 Test

This test was the same test given to the participants in Experiment One. Unlike Experiment One, the test was administered to a student as soon as the student felt prepared. Prior to taking the test the screen was cleared and any notes taken during the study time were collected.

4.3.7 Experiment Two: Procedures

The experiment took place in a closed lab with participants arriving at any time between 9:00 a.m. and 4:15 p.m. on Mondays, Wednesdays, and Fridays and between 2:00 p.m. and 4:15 p.m. on Tuesdays and Thursdays over a two-week period. The study period was not interrupted by the investigator, for example one student arrived at 4:11 p.m. and completed the experiment at approximately 5:15 p.m. All data gathering was performed by the author.

Potential participants were given an explanation of the responsibilities and risks regarding participation and were required to sign a Human Research Informed Consent Form. Students were allowed to ask questions concerning the experiment. To ensure anonymity, the participants placed the signed consent form in a manila envelope.

Once the informed consent form was read and signed by the student, the student was directed to select a group. This selection was accomplished by using the ace of clubs to signify group one, two of clubs to signify group two, and the three of clubs to signify group three. The first participant was allowed to pick from the three cards which were shuffled and arranged in no particular way. The next participant was allowed to select from the two remaining cards, and the third participant was given the last card. This process was then repeated for the next group of three students.

Once the group was selected, the participant was able to register using the registration screen. After registering, the student was then directed to a group of computers with the proper browser preloaded. This was done so that all screens in a given group showed similar content. Participants then signed onto the browser using the userid created during the registration process. They were given guidelines on what was expected to be learned from the reading both orally and in written form. The participants were given notepaper and were informed that they could write notes while studying but that the notes would be collected and retained by the investigator upon completion of the study period. The notepaper contained the userid of the student and in the cases where the student chose to interrupt the study period, the paper was collected and redistributed upon the student's return. Prior to taking the test, these papers were again collected and retained.

The browser has the ability to record the start and finish times of each site visited. This information was used to determine the total time the document was open and being studied by the participant.

At the end of the study period, each participant immediately took the test without notes or references of any kind and without time restrictions. As in the first experiment, no student took more than 15 minutes to complete the test. Once the participant had completed the experiment, the student was given a receipt indicating the amount of time involved in participation. The students were instructed to give their receipt to their respective instructors, after filling in their respective names. In this way, the investigator had no knowledge of the participant's identity and the instructors were able to assign the proper extra credit points to the individual.

The scores were tabulated and entered into a spreadsheet and SPSS for evaluation. All of the statistics used for this experiment used a 5% significance level.

4.3.8 Experiment Two: Results

Does reading a complete document, a document with context/keywords, or a document with keywords only improve test scores? To answer this question a one-way ANOVA was conducted, with test scores (with a possible range from zero to 20) as the dependent variable and groups (students who read the original document - no highlighting; students who read the context summary document - context with highlighted keywords; and students who read the keyword summary document - keywords only no highlighting) as the independent variable. The mean scores were highest for the context summary group (M=12.40, SD=2.37) and lowest for the keyword only group (M=10.70, SD=2.06). The mean score for the full document group (M=11.05, SD=2.42) was between the context and keyword groups (see Table 4.21). A boxplot of the test scores is shown in Figure 4.8. The analysis, as presented in Table 4.22, revealed differences approaching significance among groups ($F(2,57)=3.083$, $p=0.05$). The difference between the mean context score and the mean keyword score was not significant ($p=0.054$). Cohen's effect size ($f=3.083$, $d=.32$) indicates a medium to strong effect which suggests significance may exist given a sample size near 52 participants in each group. Data met homogeneity of variances of differences criteria (see Table 4.23).

Table 4.21 - Experiment Two: Mean and Standard Deviation of Keywords by Group

Group*	Mean	Std. Deviation
Full Doc	11.05	2.417
Context	12.40	2.371
Keyword	10.70	2.055
Total	11.38	2.366

*N=20 per group

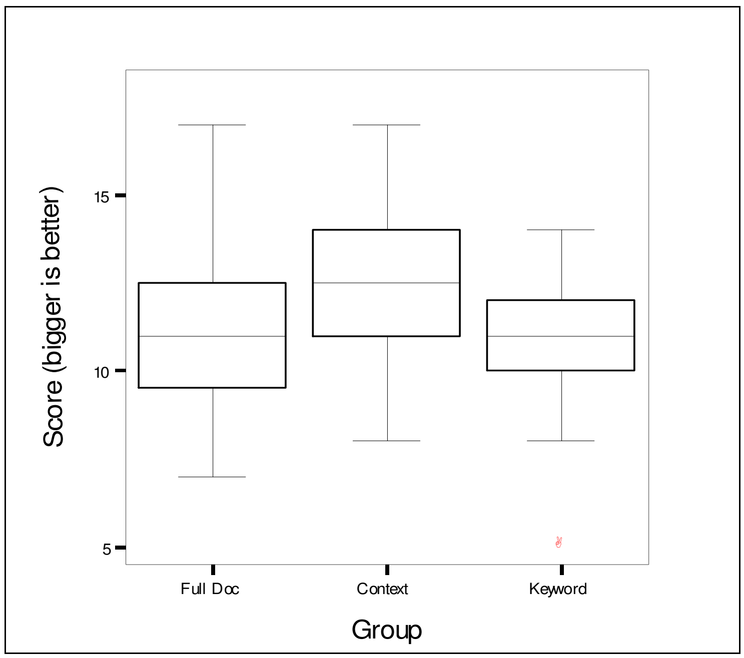


Figure 4.8 - Boxplot of Test Scores, Experiment Two

Table 4.22 - ANOVA of Test Scores by Group

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power(a)
Groupno	32.233	2	16.117	3.083	.054	.098	6.166	.572
Error	297.950	57	5.227					
Corrected Total	330.183	59						

a Computed using alpha = .05
b R Squared = .098 (Adjusted R Squared = .066)

Table 4.23 - Experiment Two Homogeneity of Variances of Differences of Test Scores

Dependent Variable: Score

F	df1	df2	Sig.
.343	2	57	.711

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a Design: Intercept+Groupno

In order to answer the second question, *"Does reading a complete document, a document with context/keywords, or a document with keywords only reduce study time?"*, preparation time was analyzed. The group that read the full document on the average (see Table 4.24) spent more time (*M*=2350.00 seconds, *SD*= 749.26) than the group that read the context/keyword summary (*M*=1787.20 seconds, *SD*=903.86). The keyword only group on the average spent the least amount of time (*M*=1684.60 seconds, *SD*=835.03). Figure 4.9 shows the boxplot of time taken for preparation for the test. A one-way ANOVA, with preparation time (in seconds) being the dependent value and three groups: students who read the original document (no highlighting), the context summary (with highlighted keywords) document, and students who read the keyword summary document (no highlighting) as the independent variable, revealed significance differences (see Table 4.25) among groups ($F(2,57)=4.06$, $p=0.02$). Homogeneity of variances of differences was assumed (see Table 4.26). Post-hoc pairwise comparisons of scores using Tukey (see Table 4.27) indicated a significant difference in preparation time between the group reading the entire document and the group reading the keyword summary.

Table 4.24 - Experiment Two: Mean and Standard Deviation of Preparation Time by Groups

Group*	Mean	Std. Deviation
Full Doc	2350.00	749.261
Context	1787.20	903.864
Keyword	1684.60	718.794
Total	1940.60	835.028

*N = 20 for all groups

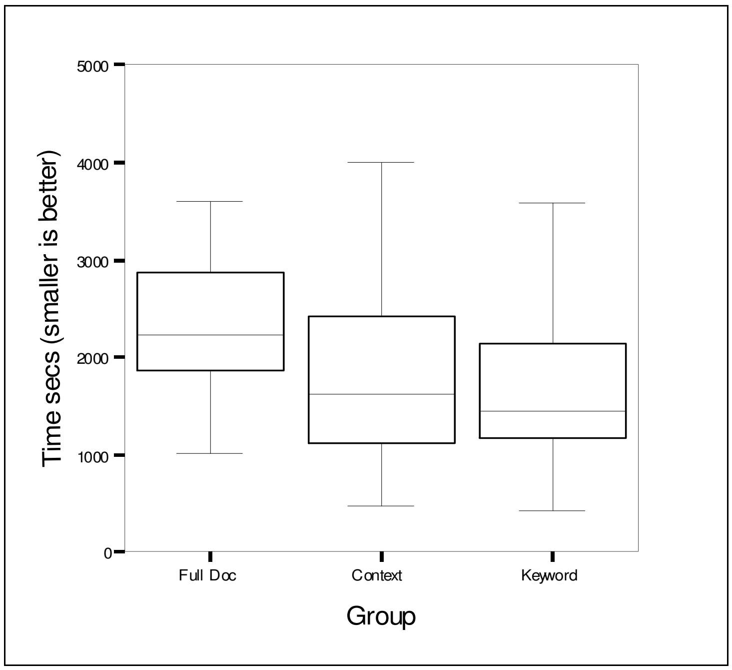


Figure 4.9 - Boxplot of Preparation Time

Table 4.25 - ANOVA to Determine Preparation Time (in Seconds) Between Groups

Dependent Variable: Timesecs

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power(a)
Groupno	5133518.400	2	2566759.200	4.063	.022	.125	8.127	.700
Error	36005526.000	57	631675.895					
Corrected Total	41139044.400	59						

a Computed using alpha = .05

b R Squared = .125 (Adjusted R Squared = .094)

Table 4.26 - Homogeneity of Variances of Differences of Preparation Time

Dependent Variable: Timesecs

F	df1	df2	Sig.
.499	2	57	.610

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a Design: Intercept+Groupno

Table 4.27 - Preparation Time, Post Hoc Test

Dependent Variable: Timesecs

	(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	Full Doc	Full Doc					
		Context	562.800	251.332	.073	-42.01	1167.61
		Keyword	665.400(*)	251.332	.028	60.59	1270.21
	Context	Full Doc	-562.800	251.332	.073	-1167.61	42.01
		Context					
		Keyword	102.600	251.332	.912	-502.21	707.41
	Keyword	Full Doc	-665.400(*)	251.332	.028	-1270.21	-60.59
		Context	-102.600	251.332	.912	-707.41	502.21
		Keyword					

* The mean difference is significant at the .05 level.

Study time and score can be combined to create a measure of efficiency using the formula:

$$Efficiency = \frac{score}{time}$$

where *score* is the number of correct answers submitted on the test and *time* is the time taken to prepare for the test. The lower the score or the greater the preparation time means a less efficient learning process. This equation was used in this experiment but may not be extended to learning efficiency overall since a preparation time approaching zero would result in an efficiency approaching infinity. In this experiment the times (and scores) were within reasonable limits.

Boxplots of the efficiency values obtained from this experiment showed a number of outliers. Logarithm_{10} of efficiency were more normally distributed hence this analysis uses $\log_{10}$(efficiency).

The final question, "*Is study time and test performance together enhanced by reading a complete document, a document with context/keywords, or a document with keywords only?",* was answered using a one-way ANOVA with efficiency, normalized using logarithm_{10}, being the dependent variable and the groups as the independent variable. The context/keyword summary group on the average had the highest efficiency score (*M*=3.33, *SD*=0.58), slightly higher than the average efficiency (*M*=3.20, *SD*=0.36) of the keyword only group (see Table 4.28). The one-way ANOVA revealed significance differences among groups ($F(2,57)=5.49$, $p=0.01$) as shown in Table 4.29. Both the context/keyword group and the keyword only group had significantly higher mean efficiency scores than the full document group (*M*=2.86, *SD*=0.41). The boxplot in Figure 4.10 shows the distribution of efficiency among the three groups. Follow-up testing was done to evaluate pairwise differences among the means. The homogeneity of variances of differences was not present (see Table 4.30), however, the group size is equal for all three groups, *N*=20. Post-hoc pairwise comparisons of scores indicated a significant difference in efficiency between the full document group and the context/keyword summary group, and the full document group and the keyword summary group (see Table 4.31). There was not a significant difference in efficiency between the context summary group and the keyword summary group.

Table 4.28 - Mean and Standard Deviation Efficiency by Group

Dependent Variable: Efficiency (log_{10})

Group*	Mean	Std. Deviation
Full Doc	2.8623	.41234
Context	3.3313	.58365
Keyword	3.2006	.36018
Total	3.1314	.49593

*N=20 for all groups

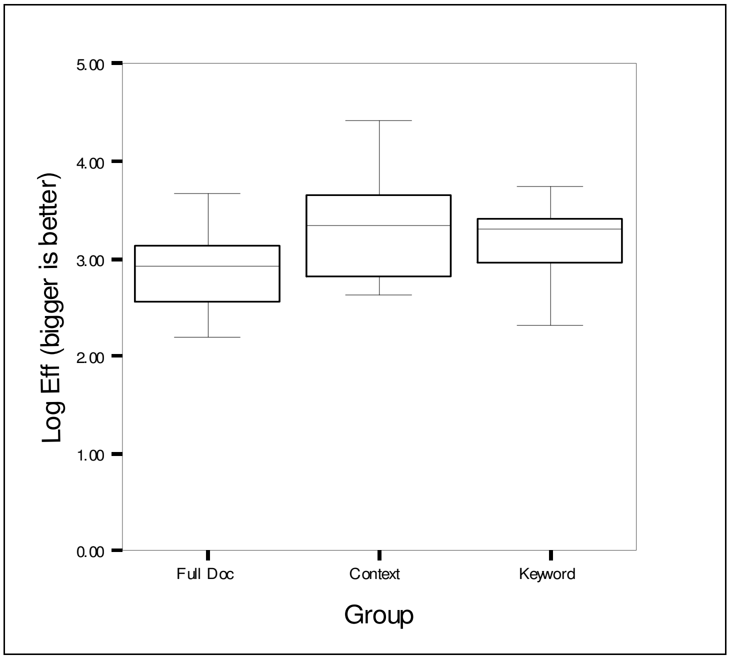


Figure 4.10 - Boxplot of Efficiency (log10)

Table 4.29 - ANOVA to Determine Efficiency (log10) Between Groups

Dependent Variable: Efficiency (log_{10})

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power(a)
Groupno	2.343	2	1.172	5.489	.007	.161	10.977	.832
Error	12.168	57	.213					
Corrected Total	14.511	59						

a Computed using alpha = .05

b R Squared = .161 (Adjusted R Squared = .132)

Table 4.30 - Homogeneity of Variances of Differences of Efficiency

Dependent Variable: Efficiency (log_{10})

F	df1	df2	Sig.
3.589	2	57	.034

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a Design: Intercept+Groupno

Table 4.31 - Efficiency (log10), Post Hoc Test

Dependent Variable: Log Eff

	(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	Full Doc	Full Doc					
		Context	-.46901(*)	.14611	.006	-.8206	-.1174
		Keyword	-.33825	.14611	.062	-.6898	.0133
	Context	Full Doc	.46901(*)	.14611	.006	.1174	.8206
		Context					
		Keyword	.13076	.14611	.646	-.2208	.4824
	Keyword	Full Doc	.33825	.14611	.062	-.0133	.6898
		Context	-.13076	.14611	.646	-.4824	.2208
		Keyword					

* The mean difference is significant at the .05 level.

4.3.9 Experiment Two: Summary

In summary, the second experiment showed the participants who viewed the context/context summary had the highest test score average of the three groups. The average test score of the context/keyword summary approached significance over the average test scores from the students viewing only a keyword summary.

The full document group, on average, spent the most time preparing for the test. The average time spent preparing for the test by the keyword only group was significantly less than the average preparation time of the full document group. While the context/keyword group on the average spent less time than the average amount of time spent by the full document group, the difference was not considered statistically different.

Efficiency is defined as the test score divided by time. The context/keyword group, on average, had the highest efficiency rating of the three groups. The full document group's average efficiency rating was significantly below the average efficiency rating of both the context /keyword and the keyword only groups. There was no significant difference in the average efficiency between the context/keyword group and the keyword only group.

4.4 Experiment Three

Experiment Three examines patterns of highlighting usage over time to see how experience alters the amount of keyword and context highlighting. In order to see if people would use HighBrow voluntarily, no incentives were provided for participation in Experiment Three.

4.4.1 Experiment Three: Goals

The goals for Experiment Three are as follows:

- To show patterns of highlighting over time
- To determine if HighBrow will be used if not required.

4.4.2 Experiment Three: Questions

The following questions are addressed by Experiment Three:

- With use, will the amount of highlighting change? If the users highlighted only keywords in the past will they begin to highlight larger portions of the text to preserve context?
- Will users voluntarily use the context browser?

4.4.3 Experiment Three: Method

In this experiment, all participants were given the MidBrow version of HighBrow, allowing them to highlight context and keywords. Over a five-week period they were given nine online documents with topics ranging from previously covered material to newly presented material

4.4.4 Experiment Three: Participants

The participants in this experiment were the students enrolled in the two sections (early morning and early afternoon) of Introduction to OOP. Because all participants were in a single group no additional selection criteria was used.

Participation was voluntary; no extra credit was awarded for participation. To provide an incentive to participate, material closely related to the topics covered in the course were used. The material may be classified as study aids.

Twenty-two students used HighBrow, however only nine actually highlighted and of these nine, only two highlighted over a period of days. Toward the latter part of the term, students elected to skip this option of the course, similar to experiences in earlier experiments (Marshall & Brush, 2002).

4.4.5 Experiment Three: Materials

The participants were given a set of nine documents extracted from Sun Microsystems online Java Tutorials (*The java tutorials.*2006) ranging in size from 359 to 8,679 words listed in Table 4.32. The documents were chosen because they contained ancillary material for the course.

4.4.6 Experiment Three: Instruments

For this experiment MidBrow was the only instrument used. Students were required to download the new version of Highbrow so that the new homepage would be shown and also to track them as participants in Experiment Three. Registration was not required as the students had already registered for Experiment One and their demographic information should have been unchanged.

Table 4.32 - Documents Viewed and/or Highlighted In Experiment Three

Title	Contents	Total words	Date added	Visitors	High-Lighters
Classes	Includes the format for defining a class with naming conventions and modifiers.	359	9/23	11	5
Variables	Good in depth discussion of variables, including the reserved word list.	2669	9/23	9	5
Operators	Includes many examples of operator usage.	1929	9/23	6	2
Methods	Answers many questions dealing with method declaration, passing and receiving data.	2343	9/23	10	2
Selection Statements	Includes many examples of if and switch code.	1012	9/23	10	1
Repetition Statements	Includes many examples of while and for code.	1599	9/23	10	3
Arrays	A nice introduction to arrays in Java.	1097	9/28	10	3
Strings	A comprehensive introduction to the String class (with converting to and from Strings, String extracting, String comparisons, etc.	3386	10/09	8	2
Classes (cont)	An in depth look at classes and objects with methods and properties.	8679	10/19	8	0
Objects and Inheritance	A nice discussion of objects and inheritance, including interfaces and packages	1685	10/23	9	1

As in Experiment One, to ensure ecological validity, the students were allowed to study the document at the time and place of their choosing.

4.4.7 Experiment Three: Procedures

All data gathering was performed by the author via HighBrow. The students were given a link to the proper installation procedures both orally and via e-mail.

Once installed and executed, the browser opened to a homepage that contained the following instructions:

> "*The following tutorials are extracts from the Java Tutorials provided by Sun Microsystems that are organized to support the lecture material in this course.*

Usage is voluntary and is subject to the informed consent rules. Students not wishing to use these tutorials may use the tutorial at the Java Sun site. Highlighting will be at your discretion.

Please visit this site often as updates will be made as the content increases."

The topics were added in synchronization with the course topics over a 30-day period, however at the beginning of the experiment several topics that were covered in earlier classes were included as study aids and for continuity. The students were verbally informed that new documents were available as new documents were added. HighBrow does record a history of mouse clicks, recording events with the userid, date and time, function, and when appropriate the start and end location of the highlight. Since this study period was not in a controlled environment, timing was not considered to be accurate (e.g., a student could open the site and be interrupted by a phone call, dinner, etc.), however the event, size and location were considered important.

4.4.8 Experiment Three: Results

Figure 4.11 shows the number of participants visiting the documents over the experiment period. The initial set of documents was released on September 23, followed by additional documents which were released on September 28, October 9, October 19, and October 23. A test on the material was given on November 2 and a third test, which also included some of the material, was given December 7.

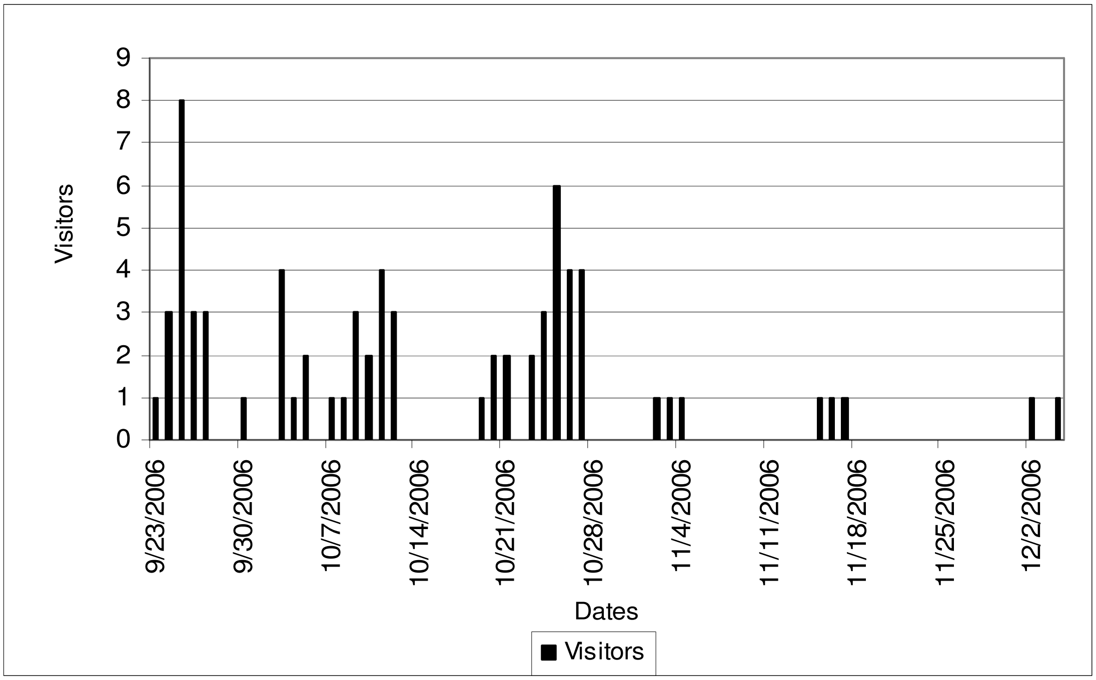


Figure 4.11 - Experiment Three Number of Visitors by Date

Figure 4.12 shows the date and users who highlighted documents over the same period of time. Only two participants highlighted over a period of time and one of those participants highlighted over a period of four consecutive days (9/24-9/27). The other participant highlighted over a period beginning 9/26 and ending 10/25. Toward the latter part of the term, students elected to skip this option of the course. No reasons were sought and none were given. Because of the lack of participation, further analysis was deemed inappropriate.

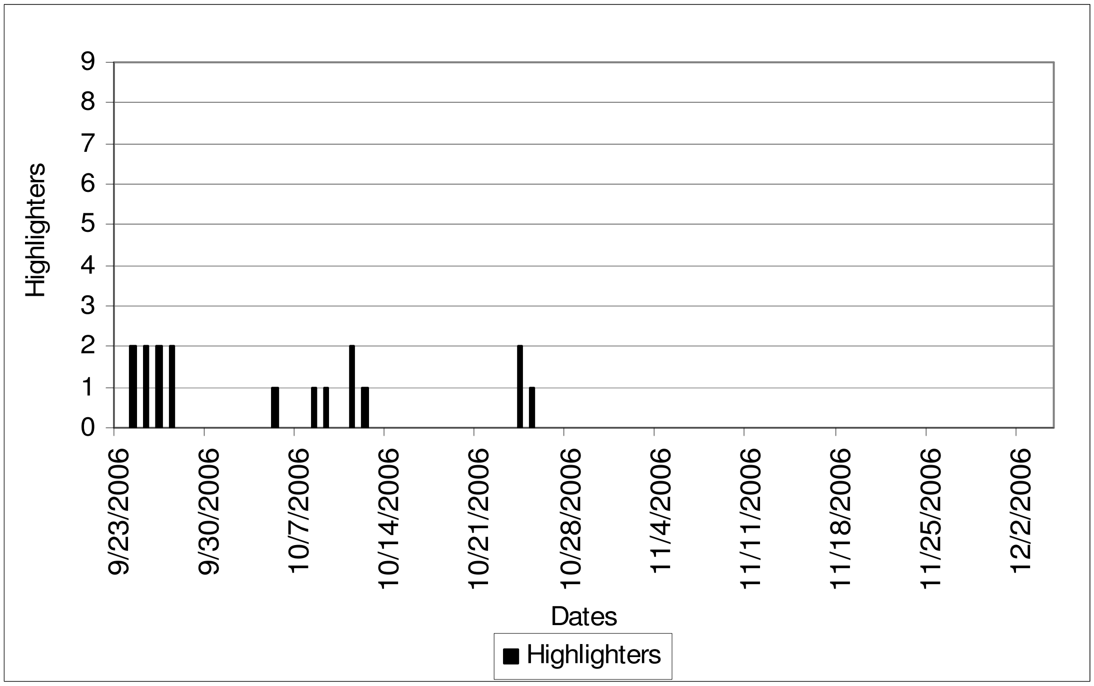


Figure 4.12 - Experiment Three Number of Highlighters by Date

4.4.9 Experiment Three: Summary

Because of the lack of participation, the results from this experiment are considered anecdotal. Table 4.33 shows the highlighting dates and the number of context and keywords highlighted on that particular date by participant.

Table 4.33 - Highlighting Dates

Date Highlighted	Participants								
	P03	P05	P07	P09	P14	P15	P17	P19	P22
	con/key	con/key	con/key	con/key	con/key	con/key	con/key	con/key	con/key
9/24/2006	160/14							9/2	
9/25/2006				0/1				329/30	
9/26/2006						52/13		150/12	
9/27/2006							1888/37	0/5	
10/5/2006					880/49				
10/8/2006			21/9						
10/9/2006					253/9				
10/11/2006		0/4				63/2			
10/12/2006									0/15
10/24/2006			0/20			128/15			
10/25/2006						16/1			

4.5 Summary of Experiments

The experiments showed that annotators with context and keywords on the average have higher test scores than annotators with keywords only. Readers of context summaries have higher average test scores than readers reading the full document or readers reading a keyword summary. Readers of context summaries spend less time on average preparing for a test than readers reading the full document, but more than readers reading a keyword summary. Readers of context summaries on the average are significantly more efficient than readers reading the full document and are more efficient than readers reading a keyword summary.

With respect to keyword size the amount of keywords highlighted by the context/keyword group on average was significantly less than the people who highlighted keywords from the keyword only group. The average amount of additional time, using the difference between the time a keyword was highlighted and the time the context was highlighted (or vice versa) was 19.1 seconds. The median time between events was nine seconds.

Chapter Five discusses the experiments and their results. The usability survey results are examined in more depth.

Chapter Five:

Discussion

The experimental results from Chapter Four indicate HighBrow does show promise as a cognitive aid for annotators of Web documents. The following sections provide discussions regarding the three experiments.

5.1 Experiment One

The test scores in Experiment One for the context/keyword group were not considered to be statistically significantly higher than the test scores from the keyword only group. The range of test scores for the context/keyword group was from 7 to18 and the keyword only group's test scores ranged from 6 to 15. The test score results are encouraging but not definitive. It was clear in Experiment One, however, that context/keyword highlighting did not lower mean test scores compared to keyword only annotation.

In Experiment One people who highlighted keywords and context spent a mean time of 19.1 seconds (median 9 seconds) between highlighting the keyword and the context or vice versa (eliminating duration times greater than three minutes). Since Experiment One was not conducted in a controlled environment, it is not certain how the preparation time was actually spent; therefore for this study, unlike Experiment Two, it is not appropriate to assign an efficiency rating using score with respect to preparation time. The difference in time to highlight context and keyword does provide a measure of the time it takes to complete the highlighting actions using HighBrow, but cannot account for the additional time spent thinking about the content, which may have helped to improve scores or interruptions, which may also have had an influence on scores.

One hypothesis in this study is: *The number of keywords highlighted in the context/keyword highlighting group would be less than the number of keywords highlighted in the keyword only group.* The difference in number of words in context versus the number of keywords may be interpreted as: more words aid interpretation and fewer words help to signal importance, as noted by Marshall (1997). Experiment One demonstrated the mean number of highlighted keywords were significantly smaller for the context/keyword participants than for the keyword only group. The size difference may be interpreted as the context/keyword group was better able to signal importance with a smaller number of keywords than the keyword only group, which may have been trying to compromise between keyword and context in a single annotation scheme.

Experiment One also showed the mean of the number of context words highlighted by the context/keyword group was larger than the mean of the number of the keywords highlighted in the keyword only group. In this case, the number of context words versus keywords may be interpreted as an indication that the participants in the context/keyword group were less constrained by the loss of importance when highlighting large passages, and were free to highlight whatever was necessary to aid interpretation. Wade and Thrathen's (1989) study suggested it is importance not the annotations that make a difference in learning.

A survey was given to the participants to determine the usability of Highbrow and to see which capabilities of HighBrow were liked. Overall, the responses to the survey indicated a very positive experience using HighBrow and that it was well liked.

An interesting response came from the item: *Context highlighting was beneficial*. Both groups had the ability to produce a context summary and both groups were allowed to highlight as much (or as little) as they wished. The responses to this question were divided sharply as shown in Table 5.1 and Table 5.2.

Table 5.1 - Context/Keyword Group, Context Was Beneficial

Response	Frequency	Percent
Neutral	2	10.5
Agree	6	31.6
Strongly agree	11	57.9
Total	19	100.0

Table 5.2 - Keyword Only Context Was Beneficial

Response	Frequency	Percent
n/a	11	55.0
Neutral	4	20.0
Agree	4	20.0
Strongly agree	1	5.0
Total	20	100.0

The context/keyword group's mean responses to the usability questions were all between agree and strongly agree except for the item: *Loading pages was fast*. The average answer was between neutral and agree with one response indicating not applicable (n/a) (see Table 5.3). For both the context/keyword and the keyword only group loading a page requires three actions: retrieval of the source document, retrieval of highlighting data from the database server, and loading and displaying the highlights on the page. The page and database retrieval take the most time since these retrievals require Web communications.

Table 5.3 - Context/Keyboard, Loading Web Pages Was Fast

Response	Frequency	Percent
n/a	1	5.3
Neutral	4	21.1
Agree	8	42.1
Strongly agree	6	31.6
Total	19	100.0

Test loads for the Web page used in the experiment showed that the mean time to load the page via Internet Explorer averaged approximately one second. Test loads for the same Web page using HighBrow varied depending on the number of highlight entries recorded. For example: a participant who highlighted 3,257 words averaged approximately 5 seconds to load the Web site, a participant who highlighted 6,141 words averaged approximately 10 seconds to load the Web site. The load time is not linear however as the load time for a user with 222,214 highlighted words took on average approximately 14 seconds to load the Web site. The time to show the highlights is almost instantaneous and is easily demonstrated by toggling the highlights on or off.

Loading the indices and displaying the highlights within the document takes very little time which was addressed by the response to a separate survey question, "*Once a page was loaded, highlighting was fast*". Overall, 22 respondents (56%) strongly agreed and 17 respondents (46%) agreed. The context/keyword group responses showed 13 participants (68%) strongly agreed and 6 participants (32%) agreed, while the keyword only group responses showed 9 participants (45%) strongly agreed and 11 participants (55%) agreed. The disparity in results was interesting because preliminary tests using both versions showed no difference in the time for HighBrow to highlight keywords or context, in fact the context/keyword group could have responded more negatively if they interpreted the question as the time to highlight both context and keyword instead of the individual highlighting action. None of the participants responded with neutral, disagree, strongly disagree, or not applicable.

The survey results indicated students liked the "*Ability to Highlight*" since both groups gave this capability the top rating (100% of the participants in the context/keyword group liked it and 95% of the participants in the keyword only group liked it) based on liked, disliked, or no opinion. The lowest response for likeability was for the "*Ability to copy Summary*" which only 55% of the participants in the keyword only group liked.

The survey contained an optional open-ended question: *What, if any, enhancements would you like to see incorporated in future versions of HighBrow?* Some students were concerned about the lack of color choices:

> "The ability to change the color of the highlights would be a nice enhancement."
>
> "highlighting in multiple color formats would be a nice addition. It would allow for a connection to be formed between content that is similar in nature within a document to be connected by highlight color. I would imagine this would be acheived [sic] through the context keywords (i.e by performing a match selected by the user while they are in the highlight

process). This connection would then make the connection and change the highlight color to match the previously highlighted text color within the document. This connecting of ideas through highlighting colors should make the content even easier to learn and remember. This is similar to what we commonly do today in our textbooks when we read and highlight the chapters."

Some students showed a desire to use it for class work:

> "I would like to see this available for the current material we are using in order to study for the upcoming test."
>
> "The highlighting was very useful. After highlighting, I only had to study the summary page for the test was given. It would be nice to somehow integrate into Word or with a .txt document."
>
> "Overall, I thought it was great, and I would use it for research if all browsers supported it."

All of the comments were not positive. One student commented:

> "Had trouble editing highlighted context and keywords. To [sic] easy to mess up. Often times it was easier to delete highlight and rehighlight."

This comment may have resulted from trying to see highlights within highlights, a definite feedback concern in the HighBrow prototype. If the user highlights keywords first, then the additional context will provide feedback, however if the student highlights context first, which appears in bright yellow as it is considered a keyword, and the user then tries to highlight keywords, HighBrow does not provide feedback. The user must remember where the highlight started and estimate using the mouse pointer, where the highlight ends. Because of this known feedback problem, the tutorial urged users to highlight keywords first.

Another participant noted,

> "The Document Text needs to be larger or you need the ability to make the text larger. Sometimes it was hard to read because of small text."

In its existing prototype state HighBrow does now allow resizing text. There is no reason why future versions could not include text resizing capability. Text resizing would not affect the presentation of highlights as the highlights are overlaid using the location of the actual text (large or small) as the anchoring point.

The participants in Experiment One's context/keyword group were new to context highlighting. I believe as annotators gain experience with this new study technique and receive feedback through test results, they may improve their ability to highlight keywords and context more accurately and efficiently.

5.2 Experiment Two

Experiment Two was intended to show the benefits of using a context/keyword summary, produced from the context/keywords highlighted by the participants from the context/keyword group in Experiment One, versus a keyword summary, produced from the keywords highlighted by the participants from the keyword only group in Experiment One, versus the full document.

One hypothesis from Experiment Two stated: "*Test performance by persons reading a context/keyword summary would be better than the test performance of persons* *reading the entire document and persons reading a keyword only summary.*" This hypothesis was based on the idea that that the context/keyword summary would contain only the relevant passages of the original document with the important words highlighted. The reader would not be distracted by words that were not relevant. The context/keyword summary would also be more beneficial than the keyword only summary, since the keyword only summary consisted mostly of tacit annotations.

A major concern in this study was if inappropriate highlighting was made by the participants in Experiment One, then this would have a negative impact on the test performance. Silvers, V. and Kreiner, D. (1997) study showed while appropriate highlighting had little effect on recall of readers reading previously annotated documents, inappropriate highlighting had a negative effect on recall. Silvers also suggested students must do the highlighting themselves in order to have any effect. The participants in Experiment One were given instructions on how to prepare for the test and were told concepts were important and actual percentages were of no concern. Despite these instructions some participants chose to highlight percentages rather than concepts. Perhaps as annotators become more proficient in highlighting, through experience, more appropriate highlighting, than the highlighting that was done in Experiment One, would create a greater benefit for both annotators and readers.

Does highlighting with context improve test scores? The hypothesis was that readers of context/keyword summaries would obtain better scores on the test. Results from Experiment Two with respect to test scores were encouraging but not conclusive. On the average, the readers of the context/keyword summary had 12% (M=12.40 versus M=11.05) higher test scores than the readers of the full document and 16% higher test scores (*M*=12.40 versus *M*=10.70) than the readers of the keyword summary; however the results lacked statistical significance.

Does reading a complete document, a document with context/keywords, or a document with keywords only reduce study time? Another hypothesis for Experiment Two was the preparation time for readers of the context/keyword summary would be less than the amount of preparation time for readers of the full document, since context/keyword summary would contain less material to read than the full document. Experiment Two revealed that there was a significant reduction in the mean preparation time between the readers of the full document and the readers of the keyword only summary document. There was a 24% reduction in the mean preparation time for the participants reading the full document (*M*=2350 seconds) and for the participants reading the context/keyword summary document (*M*=1787 seconds). However, this difference was not considered statistically significant.

Is study time and test performance together enhanced by reading a complete document, a document with context/keywords, or a document with keywords only? In this dissertation, we will define efficiency as test performance per unit of study time. Despite the fact that context highlighting is in its infant stages, the efficiency for readers reading the context summary document have on the average shown a significant improvement (14%) over the efficiency of readers reading the full document. With respect to efficiency among readers reading the keyword only summary document and the readers reading the full document, on the average the keyword document group was considered to be more efficient than the full document group, however it was not considered to be statistical significant. The keyword group took significantly less time to read the keyword summary than the full document group took to read the full document but the scores for the keyword summary group were lower than the scores from the full document group, which is not necessarily a desirable outcome. It should be noted, on the average, the context/keyword group scored higher and took less time to prepare than the full document group.

5.3 Experiment Three

The intent of Experiment Three was to see if, over time, the patterns of highlighting would change with the user creating a clearer break between keyword and context and thereby reducing the keyword size and possibly increasing the context size.

Experiment Three was most disappointing. Twenty-two students from the Introduction to OOP class (both sections) chose to participate, however only nine chose to highlight. Only two students highlighted over time, while the remaining seven appeared to wait and highlight all at once. Of the two who highlighted over a period of time, one highlighted within four consecutive days and the other highlighted over four weeks. The person who highlighted over the four weeks did in fact highlight fewer keywords and more context words but no clear trend was evident.

The documents provided for highlighting were in support of content for the course; however, many students chose not to participate. The students were obtaining similar information through lecture, assignments, their own text, etc., and perhaps were overwhelmed by the different options for receiving information. One student volunteered he would simply print the documents out so he could have a copy of the information long after the experiment was over. Attrition was very high (approximately 66%) in this introductory course and may have also been a factor in lack of overall participation. The lack of participants in a longitudinal study is not unique; Marshall and Brush (2002) noted a drop in participation as students opted to skip out of the reading and annotation experiments. Despite an extensive search, no evidence of longitudinal studies using annotation was found.

5.4 Summary

Context highlighting is a novel approach to highlighting and summarizing documents. In it infancy, it has shown promise for both annotators and readers. With any new idea, as it matures, we discover new areas to develop. Chapter Six discusses possible areas for further research into the benefits and improvements of context highlighting.

Chapter Six:

Contribution and Future Work

6.1 Contribution

This dissertation makes the following contributions to the areas of HCI and cognition summarized below:

- Human Computer Interaction: Context highlighting is a new approach that provides a simple unified technique for highlighting keywords and the context surrounding them. The annotator is no longer required to change roles from reader to writer and back, nor is the annotator required to change tools (e.g., marker to pen and back) in order to create keyword highlighting and context summaries. The process of creating useful readable summaries, with emphasis on keywords, without requiring typing is now possible. The prototype context browser, HighBrow was easy to use and well liked by the users.
- Cognition: There was no hint of negative test performance by the annotators using context/keyword highlighting than the keyword only annotators; in fact the mean of the scores were marginally higher than the scores obtained by the keyword only annotators. Readers who read the summaries created by the annotators were more efficient, with respect to test scores and preparation time, than readers of the full document. The study also revealed that readers of keyword only summaries, while being more efficient than readers reading the

whole document, were slightly less efficient than the readers of the context/keyword only document. The high efficiency of readers of the keyword only summary was a result of reduced preparation time but was negatively impacted by lower test performance. The reader of the context/keyword summary benefited from less preparation time than the reader of the full document and better test performance than both the test scores from the readers of the full and keyword only documents.

The experiments conducted using HighBrow, while promising; show there is more to be done with respect in the areas of cognition, HCI, and software engineering.

6.2 Future Work: Cognition Areas

Experiment One was conducted with the students using HighBrow in an environment that was similar to a normal study environment. The participants could eat, listen to music, radio, TV, etc., be interrupted by family, friends co-workers etc. all of which are possible conditions during actual reading and preparation. In order to concentrate on the efficiency of the context highlighting process, a controlled experiment similar to Experiment Two, should be conducted to determine if there is a cost, with respect to additional time spent highlighting context, that context highlighting may incur over keyword only highlighting.

A longitudinal study should also be conducted in a controlled environment to ensure usage over time, rather than a last minute one-time only preparation, so participants may be able to mature using context highlighting. The longitudinal study could also test for short and long term recall of key concepts. A longitudinal study would have to have numerous homogenous documents of sufficient size to require highlighting that would be of interest to the participants. The longitudinal study would determine changes in highlighting patterns over time and also determine if context and keyword highlighting can be improved with practice.

It would be interesting to see if “experts” using context highlighting would produce more useful summary documents for readers than the average annotator. By highlighting only the “appropriate” content, the important data would be included in the summaries and would not mislead or confuse the reader. Will there be improved test scores, reduced preparation time, and higher efficiency ratings resulting from summaries produced by experts over the summaries produced from the average reader? What significant differences are there in summaries produced by experts and summaries produced by the average user?

Educators and psychologists may be able to determine learning disabilities by comparing summaries of keywords highlighted by students versus keywords created by “experts”. Wade and Trathen (1989) said, “Lower ability students may differ from higher ability students by reading either with no criteria of importance in mind or with different criteria. Or, they may be as sensitive to importance as higher ability students but less capable or consistent in using strategies to learn what they identified as important.” Can context highlighting, by signaling importance through more focused keywords, result in more accuracy in determining importance?

6.3 Future Work: Human Computer Interaction

Refinding earlier accessed pages represent a significant problem for many users of the World Wide Web. Studies could be conducted to determine the benefits of using keyword highlights, rather than context, as a more specific way to index content already seen than the traditional “favorites” or “history” tools provided by existing browsers. Web pages are more dynamic, the highlights will also help to determine if the material sought is still present which may result in early termination of a search, rather than continuing with “I know it is around here someplace.”

HighBrow was never intended to be a finished product, it was meant to be a prototype instrument to test the concept of context highlighting. HighBrow was successful for its intended purpose, however in daily usage it would need to be a fully functional browser much like Mozilla’s FireFox, Microsoft’s Internet Explorer, etc. A challenge here is to maintain the simplicity of HighBrow’s annotation and summary capability which help invite usage. As many other annotation developers have noted, it must be simple.

The remaining future work in this section may be better described as HCI enhancements or features:

- HighBrow used bright yellow for keywords and light yellow for context and the source document, by design, had black text and white backgrounds ensuring contrast. The Web provides a variety of backgrounds (including pictures) and text colors. Eliminating the requirement to choose a color for highlighting makes it that much easier. A good enhancement would be to have the software automatically select the highlighting colors based on the source document's background and text colors, eliminating the need for the user to have to select the colors for highlighting.

- The context summary is very rudimentary, showing the title of the original document, if given, highlighted keywords, and the context. The user may annotate the summary using typed text, however the typed text is not saved with the document (it can, however, be saved as a separate file or printed). Providing a way to save the added notation is an enhancement that may be worth pursuing.

6.4 Future Work: Software Engineering

Initial usage of HighBrow has been brief and controlled, however if context highlighting is successful, then there are some basic software engineering concerns which must be addressed.

The list of highlights may grow at a tremendous rate, creating a problem for the system to maintain a list of all highlights. The problem will not be at the database server level as database servers are capable of handling the data. The problem will be in communicating and storing the annotations on the client machine as was demonstrated in the time taken to load the page. Possible solutions may be for the user to have annotation maintenance control panels to eliminate unwanted or unused annotations; however using a maintenance control panel violates the HCI simplicity model that is desired. Handling the volume of annotations between database server and client software, while maintaining the freedom to use any computer anywhere, will remain an open problem.

Security issues regarding the transmitting of sensitive selected material may also be an issue. Should the highlight metadata be encrypted and if so, how should it be encrypted?

6.5 Summary

Context highlighting may be a powerful tool for users of the Web to help summarize and retain information; however, context highlighting has opened up a new set of questions and challenges for psychologists, educators, and software engineers alike. The results from this dissertation indicate the potential benefits of context highlighting make these questions and challenges worthy of pursuit.

References

Annotea project. (2005). Retrieved May 8, 2006, from http://www.w3.org/2001/Annotea/.

Benford, S., Schnädelbach, H., Koleva, B., Anastasi, R., Greenhalgh, C., & Rodden, T., et al. (2005). Expected, sensed, and desired: A framework for designing sensing-based interaction. [Electronic version]. *ACM Trans.Comput.-Hum.Interact., 12*(1), 3-30.

Bottoni, P., Civica, R., Levialdi, S., Orso, L., Panizzi, E., & Trinchese, R. (2005). Digital library content annotation with the MADCOW system. Paper presented at the *Proceedings of the 7th International Workshop of EU Network of Excellence on Audio-Visual Content and Information Visualization in Digital Libraries,* Cortona, Italy. 111-116. Retrieved June 12, 2006, from http://www.telecom.gouv.fr/programmes/eten/madcow.pdf.

Brennan, S., Winograd, P. N., Bridge, C. A., & Hiebert, E. H. (1986). A comparison of observer reports and self-reports of study practices used by college students. *National Reading Conference Year Book, , 35* 353-358.

Brown, P. J., & Brown, H. (2004). Integrating reading and writing of documents. [Electronic version]. *Journal of Digital Information, 5*(1)Retrieved May 2, 2006.

Bush, V. (1945). *As we may think.* Retrieved July 12, 2006, from http://ccat.sas.upenn.edu/~jod/texts/vannevar.bush.html.

Cousins, S., Baldonado, M., & Paepcke, A. (2000). *A systems view of annotations*. Retrieved August 18, 2006, from http://www.informatics.indiana.edu/dgroth/Research/Projects/Annotation/Readings/Systems_View_Annotations.pdf.

Crystal, M. R., Kubala, F., & MacIntyre, R. (1999). Studies in data annotation effectiveness. Paper presented at the *Proceedings of the DARPA Broadcast News Workshop,* Herndon, Virginia. Retrieved July 20, 2006, from http://www.nist.gov/speech/publications/darpa99/index.htm.

Definition of annotation - merriam-webster online dictionary. Retrieved December 10, 2006, from http://mw1.merriam-webster.com/dictionary/annotation.

Denoue, L., & Vignollet, L. (2000). An annotation tool for web browsers and its applications to information retrieval. Paper presented at the *RIAO2000 (Recherche d'Information Assistée Par Ordinateur),* Paris, France. Retrieved June 6, 2006, from http://www.univ-savoie.fr/labos/syscom/Laurent.Denoue/riao2000.doc.

Denoue, L. (2005). *YAWAS.* Retrieved July 12, 2006, from http://www.fxpal.com/people/denoue/yawas/.

Effect of computers on workplace stress, job security, and work interest in canada. (2005). Retrieved July 20, 2005, from http://www11.hrsdc.gc.ca/en/cs/sp/hrsdc/arb/publications/research/2002-000146/page00.shtml.

Fu, X., Ciszek, T., Marchionini, G., & Solomon, P. (2005). Annotating the web: An exploratory study of web users' needs for personal annotation tools. Paper presented at the *68th Annual Meeting of the American Society for Information Science and Technology,* Charlotte (US). Retrieved July 18, 2006, from http://eprints.rclis.org/archive/00005095/.

Gibbons, S., Peters, T. & Bryan, R. (2003). *E-book functionality white paper.* Retrieved October 12, 2006, from http://www.lib.rochester.edu/main/ebooks/ebookwg/white.pdf.

Golovchinsky, G., Price, M. N., & Schilit, B. N. (1999). From reading to retrieval: Freeform ink annotations as queries. *SIGIR '99: Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval,* Berkeley, California, United States. 19-25. from http://doi.acm.org/10.1145/312624.312637.

Herman, I. (2007). *Semantic web activity statement.* Retrieved February 14, 2007, from http://www.w3.org/2001/sw/Activity.

Herman, I., Swick, R. & Brickley, D. (2007). *Resource description framework (RDF).* Retrieved February 14, 2007, from http://www.w3.org/RDF/.

Hershberger, W. (1964). Self-evaluational responding and typographical cueing: Techniques for programing self-instructional reading materials. *Journal of educational psychology, 55*(5), 288-296.

Huynh, D., Mazzocchi, S. & Lee, R. (2007). *Piggy bank - SIMILE.* Retrieved February 8, 2007, from http://simile.mit.edu/wiki/Piggy_Bank.

i-lighter the yellow marker on the web. (2006). Retrieved January 23, 2007, from http://www.i-lighter.com/.

The java tutorials. (2006). Retrieved September 20, 2006, from http://java.sun.com/docs/books/tutorial/.

Levy, D. M. (1997). I read the news today, oh boy: Reading and attention in digital libraries. Paper presented at the *DL '97: Proceedings of the Second ACM International Conference on Digital Libraries,* Philadelphia, Pennsylvania, United States. 202-211. from http://doi.acm.org/10.1145/263690.263817.

Lorch, R. F. J., Pugzles-Lorch, E., & Klusewitz, M. A. (1995). Effects of typographical cues on reading and recall of text. [Electronic version]. *Contemporary educational psychology, 20*(1), 51-64. Retrieved Sept 24, 2006, from ERIC database.

Marshall, C. C. (1997). Annotation: From paper books to the digital library. *DL '97: Proceedings of the Second ACM International Conference on Digital Libraries,* Philadelphia, Pennsylvania, United States. 131-140. from http://doi.acm.org/10.1145/263690.263806.

Marshall, C. C. (1998). Toward an ecology of hypertext annotation. *HYPERTEXT '98: Proceedings of the Ninth ACM Conference on Hypertext and Hypermedia : Links, Objects, Time and Space---Structure in Hypermedia Systems,* Pittsburgh, Pennsylvania, United States. 40-49. from http://doi.acm.org/10.1145/276627.276632.

Marshall, C. C., & Brush, A. J. B. (2002). From personal to shared annotations. *CHI '02: CHI '02 Extended Abstracts on Human Factors in Computing Systems,* Minneapolis, Minnesota, USA. 812-813. from http://doi.acm.org/10.1145/506443.506610.

Nist, S. L., & Hogrebe, M. C. (1987). The role of underlining and annotating in remembering textual information. [Electronic version]. *Reading Research and Instruction, 27*, 12-25. from EDUFT database.

Obendorf, H. (2003). Simplifying annotation support for real-world-settings: A comparative study of active reading. *HYPERTEXT '03: Proceedings of the Fourteenth ACM Conference on Hypertext and Hypermedia,* Nottingham, UK. 120-121. from http://doi.acm.org/10.1145/900051.900076.

Olsen, D. R., Taufer, T., & Fails, J. A. (2004). ScreenCrayons: Annotating anything. *UIST '04: Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology,* Santa Fe, NM, USA. 165-174. from http://doi.acm.org.dax.lib.unf.edu/10.1145/1029632.1029663.

Ostler, T. (1999). Information highlighting. Paper presented at the *1999 IEEE International Conference on Information Visualization,* London, UK. 528-534. Retrieved June 12, 2006, from http://ieeexplore.ieee.org/xpl/RecentCon.jsp?punumber=6353.

Ovsiannikov, I. A., Arbib, M. A., & McNeill, T. H. (1999). Annotation technology. *International Journal of Human-Computers Studies, 50*(4), 329-362.

Peterson, S. E. (1992). The cognitive functions of underlining as a study technique. *Reading Research and Instruction, 31*(2), 49-56.

Phelps, T. A., & Wilensky, R. (1996). Toward active, extensible, networked documents: Multivalent architecture and applications. *DL '96: Proceedings of the First ACM International Conference on Digital Libraries,* Bethesda, Maryland, United States. 100-108. from http://doi.acm.org/10.1145/226931.226951.

Phelps, T. A., & Wilensky, R. (1997). Multivalent annotations. *ECDL '97: Proceedings of the First European Conference on Research and Advanced Technology for Digital Libraries,* 287-303.

Quan, D. A., & Karger, R. (2004). How to make a semantic web browser. *WWW '04: Proceedings of the 13th International Conference on World Wide Web,* New York, NY, USA. 255-265. from http://doi.acm.org/10.1145/988672.988707.

Roscheisen, M., Mogensen, C., & Winograd, T. (1997). *Shared web annotations as a platform for third-party value-added, information providers: Architecture, protocols, and usage examples*. Stanford, CA, USA: Stanford University.

Saltzer, J. H., Reed, D. P., & Clark, D. D. (1984). End-to-end arguments in system design. [Electronic version]. *ACM Trans.Comput.Syst., 2*(4), 277-288.

Schilit, B. N., Golovchinsky, G., & Price, M. N. (1998). Beyond paper: Supporting active reading with free form digital ink annotations. *CHI '98: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* Los Angeles, California, United States. 249-256. from http://doi.acm.org/10.1145/274644.274680.

Shilman, M., & Wei, Z. (2004). Recognizing freeform digital ink annotations. Paper presented at the *6th International Workshop, DAS 2004,* Florence, Italy. 322-331. from http://www.informatik.uni-trier.de/~ley/db/conf/das/das2004.html#ShilmanW04.

Silvers, V., & Kreiner, D. (1997). The effects of pre-existing inappropriate highlighting on reading comprehension. [Electronic version]. *Reading Research And Instruction, 36*, 217-223. Retrieved August 7, 2006.

Turney, P. (1999). *Learning to extract keyphrases from text* No. NRC/ERB-1057)National Research Council of Canada. Retrieved September 15, 2006, from https://iit-iti.nrc-cnrc.gc.ca/iit-publications-iti/docs/NRC-41622.pdf.

Vatton, I. (2006). *Amaya screenshots.* Retrieved February 16, 2007, from http://www.w3.org/Amaya/screenshots/Overview.html.

Vatton, I. (2007). *Amaya home page.* Retrieved February 8, 2007, from http://www.w3.org/Amaya/.

Wade, S. E., & Trathen, W. (1989). Effect of self-selected study methods on learning. [Electronic version]. *Journal of educational psychology, 81*(1), 40-47. Retrieved July 12, 2006, from ERIC database.

XANADU® ARCHIVE PAGE. Retrieved June 12, 2006, from http://www.xanadu.com/XUarchive/.

Yee, K. (2002). CritLink: Advanced hyperlinks enable public annotation on the web. Demonstration Abstract *ACM Conference on Computer-Supported Co-Operative Work,* New Orleans.

Bibliography

Abrams, D., Baecker, R., & Chignell, M. (1998). Information archiving with bookmarks: Personal web space construction and organization. *CHI '98: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* Los Angeles, California, United States. 41-48. from http://doi.acm.org/10.1145/274644.274651.

Adler, M. J. (1960). *How to read a book*. New York, NY: Simon and Schuster.

Bailey, R. J. (2006). The effects of highlighting on long-term memory. Retrieved December 19, 2006, from http://clearinghouse.missouriwestern.edu/manuscripts/294.asp.

Bargeron, D., & Moscovich, T. (2003). Reflowing digital ink annotations. CHI '03: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Ft. Lauderdale, Florida, USA. 385-393. from http://doi.acm.org/10.1145/642611.642678.

Brush, A. J. B. (2002). Annotating digital documents: Anchoring, educational use, and notification. *CHI '02: CHI '02 Extended Abstracts on Human Factors in Computing Systems,* Minneapolis, Minnesota, USA. 542-543. from http://doi.acm.org/10.1145/506443.506472.

Cockburn, A., & Greenberg, S. (2000). *Issues of page representation and organisation in web browser's revisitation tools.* Retrieved June 7, 2006, from http://www.cosc.canterbury.ac.nz/andrew.cockburn/papers/issuesOzCHI.pdf.

Denoue, L., & Vignollet, L. (2000). New ways of using web annotations. Paper presented at the *9th International World Wide Web Conference,* Amsterdam. Retrieved June 8, 2006, from http://www9.org/final-posters/poster46.htm.

Edwards, P. N. (2005). *How to read a book strategies for getting the most out of non-fiction reading.* Retrieved February 1, 2007, from http://www.si.umich.edu/~pne/PDF/howtoread.pdf.

Gordon, M., & Pathak, P. (1999). Finding information on the world wide web: The retrieval effectiveness of search engines. *Inf.Process.Manage., 35*(2), 141-180.

Green, S. B., & Salkind, N. J. (2003). *Using SPSS for windows and macintosh* (3rd ed.). Upper Saddle River, New Jersey: Prentice Hall.

GVU's seventh WWW user survey: Problems using the web graphs. Retrieved September 20, 2006, from http://www.gvu.gatech.edu/user_surveys/survey-1997-04/graphs/use/Problems_Using_the_Web.html.

Hansen, F. A. (2006). Ubiquitous annotation systems: Technologies and challenges. *HYPERTEXT '06: Proceedings of the Seventeenth Conference on Hypertext and Hypermedia,* Odense, Denmark. 121-132. from http://doi.acm.org/10.1145/1149941.1149967.

Hightower, R. R., Ring, L. T., Helfman, J. I., Bederson, B. B., & Hollan, J. D. (1998). PadPrints: Graphical multiscale web histories. *UIST '98: Proceedings of the 11th Annual ACM Symposium on User Interface Software and Technology,* San Francisco, California, United States. 121-122. from http://doi.acm.org/10.1145/288392.288582.

LaLiberte, D., & Braverman, A. (1997). A protocol for scalable group and public annotations. Paper presented at the *Third International World-Wide Web Conference,* Darmstadt, Germany. Retrieved September 23, 2006, from http://www.igd.fhg.de/archive/1995_www95/papers/100/scalable-annotations.html.

Lebow, D., & Lick, D. (2002). HyLighting A new tool for distance and distributed teaching and learning. Paper presented at the *The Eighth Sloan-C International Conference,* Retrieved September 22, 2006, from http://www.sloan-c.org/conference/proceedings/2002/track6.asp.

Li, W., Vu, Q., Chang, E., Agrawal, D., Hirata, K., & Mukherjea, S., et al. (1999). PowerBookmarks: A system for personalizable web information organization, sharing, and management. *SIGMOD '99: Proceedings of the 1999 ACM SIGMOD International Conference on Management of Data,* Philadelphia, Pennsylvania, United States. 565-567. from http://doi.acm.org/10.1145/304182.304578.

Lopez, M. J. M., Rodriguez, M. d. B., & Hidalgo, J. M. G. (1999). Using and evaluating user directed summaries to improve information access. *ECDL '99: Proceedings of the Third European Conference on Research and Advanced Technology for Digital Libraries,* 198-214.

Maarek, Y. S., & Shaul, I. Z. B. (1996). Automatically organizing bookmarks per contents. Paper presented at the *Fifth International World Wide Web Conference on Computer Networks and ISDN Systems,* Paris, France. 1321-1333. from http://dx.doi.org.dax.lib.unf.edu/10.1016/0169-7552(96)00024-4.

Marais, H., & Bharat, K. (1997). Supporting cooperative and personal surfing with a desktop assistant. *UIST '97: Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology,* Banff, Alberta, Canada. 129-138. from http://doi.acm.org/10.1145/263407.263531.

Meyer, B. J. F., Brandt, D. M., & Bluth, G. J. (1980). Use of top-level structure in text: Key for reading comprehension of ninth-grade students. [Electronic version]. *Reading Research Quarterly, 16*(1), 72-103. Retrieved September 23, 2006, from SFX database.

O'Hara, K., & Sellen, A. (1997). A comparison of reading paper and on-line documents. *CHI '97: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* Atlanta, Georgia, United States. 335-342. from http://doi.acm.org/10.1145/258549.258787.

Phelps, T. A., & Wilensky, R. (2000a). *Robust hyperlinks cost just five words each.* Berkeley, CA, USA: University of California at Berkeley.

Phelps, T. A., & Wilensky, R. (2000b). Robust intra-document locations. *Proceedings of the 9th International World Wide Web Conference on Computer Networks : The International Journal of Computer and Telecommunications Netowrking,* Amsterdam, The Netherlands. 105-118. from http://dx.doi.org/10.1016/S1389-1286(00)00043-8.

Shneiderman, B. (1998). *Designing the user interface: Strategies for effective human-computer interaction* (3rd ed.)Addison Wesley Longman, Inc.

Tombros, A., & Sanderson, M. (1998). Advantages of query biased summaries in information retrieval. *SIGIR '98: Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval,* Melbourne, Australia. 2-10. from http://doi.acm.org/10.1145/290941.290947.

Vasudevan, V., & Palmer, M. (1999). On web annotations: Promises and pitfalls of current web infrastructure. *HICSS '99: Proceedings of the Thirty-Second Annual Hawaii International Conference on System Sciences-Volume 2,* 2012.

Weiss, R., Velez, B., & Sheldon, M. A. (1996). HyPursuit: A hierarchical network search engine that exploits content-link hypertext clustering. *HYPERTEXT '96: Proceedings of the Seventh ACM Conference on Hypertext,* Bethesda, Maryland, United States. 180-193. from http://doi.acm.org.dax.lib.unf.edu/10.1145/234828.234846.

Wolfe, J. L. (2000). Effects of annotations on student readers and writers. *DL '00: Proceedings of the Fifth ACM Conference on Digital Libraries,* San Antonio, Texas, United States. 19-26. from http://doi.acm.org/10.1145/336597.336620.

Xin, C., & Feenberg, A. (2002). Designing for pedagogical effectiveness: The TextWeaver. Paper presented at the *35th Annual Hawaii International Conference on System Sciences,* 1090-1099. Retrieved July 12, 2006, from http://ieeexplore.ieee.org.ezproxy.lib.usf.edu/xpl/tocresult.jsp?isnumber=21442&isYear=2002.

Zeiliger, R., Reggers, T., Baldewyns, L., & Jans, V. (1997). Facilitating web navigation : Integrated tools for active and cooperative learners. Paper presented at the *5th International Conference on Computers in Education,* Kuching, Malaysia. Retrieved June 23, 2006, from http://www.gate.cnrs.fr/~zeiliger/artWN97.doc.

Appendices

Appendix A: Full Document Used in Experiment One and Experiment Two

Excerpts reproduced from the Government of Canada, Department of Human Resources and Social Development. Information study: THE EFFECTS OF COMPUTERS ON WORKPLACE STRESS, JOB SECURITY AND WORK INTEREST IN CANADA - DECEMBER 2002

The Effects of Computers on Workplace Stress, Job Security and Work Interest in Canada - December 2002

Information and communication technologies (ICTs), especially computers, have dramatically changed the way we work and live. According to the 2000 General Social Survey (GSS 2000) of Statistics Canada, nearly six out of ten Canadian workers used a computer (personal computer, mainframe or word processor) at work, with the majority (78%) using it to perform various tasks on a daily basis (Marshall (2001)). This usage rate is up from one in two in 1993 (Morissette and Drolet (1998)) and from 39% in 1989 (Lowe (1997)).

The majority of earlier literature centers around the effects on productivity and job quality, which Rubery and Grimshaw (2001) sub-divide into three main dimensions: 1) employment relations and protection (e.g., employment opportunities, employment relations, career opportunities, job protection and collective bargaining, pay); 2) time and work autonomy (e.g., work intensity, power and autonomy, work/life balance, work relations); and 3) skills and careers (e.g., skills, job prospects). Polarized views on the effects of ICTs can be found in the literature in terms of each of these dimensions. For example, the pessimistic argues that ICTs destroy employment opportunities through automation and rationalization, reduce pay by downgrading skills and weakening workers' collective bargaining power. To the exact opposite, the optimistic hypothesizes that ICTs create jobs through developing new markets and human capital, increase pay by augmenting skills.

Many previous studies have found a positive relationship between productivity and the use of ICTs (e.g., Greenan and Mairesse (2000); Gera, Gu and Lee (1999); Brynjolfsson and Hitt (1996); Lichtenberg (1995); Siegel and Griliches (1992)). Empirical evidence on the positive association between wages and the use of ICTs is also abundant (e.g., Autor, Latz and Kruger (1998); Baldwin, Gray and Johnson (1997); Bound and Johnson (1992)). While supporting the finding that there is a positive linkage between wages and the use of computers and other advanced technologies, work by others (e.g., Morissette and Drolet (1998); Dinardo and Pischke (1997); Entorf and Kramarz (1996)) has argued that workers who use computers earn more than other employees (those who do not) not because of their computing skills per se, but rather because they are endowed with more other unobservable or unmeasureable skills.

Other work (e.g., Baldwin, Diverty and Sabourin (1995)) has shown that the adoption of computers and other new technologies is a key element to firms' success because these technologies are correlated with market share increases, productivity gains, product and delivery quality improvements, increased flexibility, production costs reduction, and so on (e.g., Baldwin and Lin (2002)).

There are many other aspects on which ICTs may have significant impacts too. For example, using a special supplement to the December 1998 Current Population Survey (CPS), Kuhn and Skuterud (2000) show that 15% of unemployed job seekers in the United States used the internet to search for jobs in 1998, so did half of all job seekers with online access from home. They further demonstrate that internet job search rates exceeded those of traditional job search methods such as services provided by private employment agencies, contacting friends and relatives, using trades unions or professional associations.

With employers fiercely competing for technological advantages by widely adopting and frequently upgrading ICTs, workers constantly find themselves being surrounded by these technologies. What impacts does the adoption of ICTs have on workers? Specifically, what psychological impacts do ICTs and the constant need to learn new computer skills may have on workers? Do they cause extra stress or worry? Are some workers affected more than others? Answers to questions such as these are important if we are going to better understand the profound impacts the ICTs revolution has caused. To quote a commentary in CQ Researcher, "The computer revolution has given the modern workplace an array of new options and improved efficiency. But far from having a calming effect on overworked employees, computerization has itself become a source of increasing psychological stress." (August 14, 1992: 703)

Further, although actual overall job stability and security changed very modestly in both Canada and the United States up to the mid-1990s, this small change in the aggregate masks rather sharp declines or rises for certain groups of workers (e.g., Neumark, Polsky and Hansen (1999), Picot, Lin and Pyper (1998), Schmidt and Svorny (1998), Picot and Lin (1997)). Does the adoption of ICTs contribute in any way to changes in job security and stability? If so, are the impacts felt uniformly across the board or diversely across different groups of workers?

In addition, ICTs have increasingly replaced humans to perform a great number of complex and challenging tasks. As a result, many processes and tasks have been automated or routinized. Has this made work more or less interesting/boring? If yes, are the impacts invariant across all workers or some workers are affected more than others?

Appendix A: (Continued)

While effects of ICTs on such areas as productivity, wages, firm performance are well researched and documented, there has been far less effort and work on job quality as measured in psychological stress, job security and work interest. The objective of this

paper is hence to add to the literature empirical evidence on effects of ITCs on dimensions as mentioned above.

Using the nationally representative survey on access to and use of information and communication technology of Statistics Canada, this paper attempts to empirically address the following specific questions: Does having to learn new computer skills cause extra stress? Do computers affect work and to what extent? To the extent that work is affected, do computers make job more or less secure? Do computers make work more or less interesting? Are the effects of computers on these measures of job quality felt invariably in the same way by all workers or differently by workers with different attributes?

The presentation of materials proceeds as the following. Section 2 briefly describes the data used for the analysis, discusses our model and explanatory variable specifications, sample restrictions, and estimation. Section 3 presents and discusses our results on the effects of computers and automated technology: 3.1 for findings on stress in the workplace; 3.2 on work being affected; 3.3 on job security change; and 3.4 on work interest change. Finally, Section 4 concludes.

2. Data, Model, Sample, and Estimation

We use data extracted from the public use microdata file of the 14th cycle of the General Social Survey of Statistics Canada, conducted from January through December 2000 (GSS 2000). The target population for this survey is all Canadians 15 years of age and older, who are not residents of the three territories (Yukon, Northwest and Nunavut) or full-time residents of institutions (e.g., the armed forces, correctional facilities, health-care institutions).

GSS 2000 is a household-based survey and has 25,090 respondents, representing approximately 24.6 million Canadians. It contains a wealth of information on access to and use of ICTs in Canada, especially computers and the internet, in the 12 months prior to the survey date. It also contains a wealth of information on respondents' personal and socio-economic characteristics.

All research questions addressed in the paper are derived from the GSS direct questioning of respondents on the effects of computers and automated technology. For notational convenience, we term "computers and automated technology" in short as "computers".

The first question we try to address is whether having to learn new computer skills causes excess worry or stress in the workplace and if so, whether the stress varies with observable demographic attributes, geographic locations, and work characteristics.

The second question we attempt to answer is whether work is affected by computers and if so, the extent to which work is affected. The survey provides four mutually exclusive

answers. We combine the "hardly" and "not at all" cases into one category and thus, the dependent variable takes on three values: one for "work being greatly affected"; another for "work being somewhat affected"; and the remaining for "work being hardly or not at all affected". The sub-sample of those who state that their work is greatly or somewhat affected by computers is further asked if computers have changed their job security and work interest.

Our third research question is thus how computers have changed job security. In the context of the survey, the dependent variable has three discrete and ordinal values: one for "job security has increased"; another for "job security has stayed the same"; and the rest for "job security has decreased".

Finally, the fourth question addressed in the paper is how computers have changed work interest. The survey provides three mutually exclusive answers: "work has become more interesting"; "no change in work interest"; and "work has become less interesting".

Canada is a large country composed of economically diverse regions. As computer use varies somewhat from one area to another (Lin and Popovic (2002a)), effects of computers are also expected to vary. Hence, geographic locations indicated by province and urban/rural area of residence are entered into the models as additional regressors.

Further, as computer usage differs substantially across a set of work characteristics, effects of computers are also expected to vary along these dimensions. Within the context of our data, these work characteristics include full-time or part-time work schedule, employee or self-employed (with or without paid help) employment type, industry, and occupation.

The final empirical samples used to estimate these equations include respondents aged 15 to 64 who were not full-time students at the time of the survey and were at work during the reference week. The sample for the stress equation consists of 7,741 observations, representing about 7.9 million workers who used computers at work. The work effect equation is modelled with the sample of 13,150 observations, representing about 13.4 million workers. The job security model is estimated on the sample of 7,744 observations, representing about 7.9 million workers who stated that their work has been greatly or somewhat affected by computers. Finally, the sample used to estimate the work interest model is made up of 7,779 observations, representing about 7.9 million workers who stated that their work has been greatly or somewhat affected by computers.

Of course, it also suffers numerous limitations. Most noticeably, the lack of information on employer characteristics (e.g., firm size by the number of employees or assets/revenues; ownership (Canadian vs foreign); human resources management practices such as compensation pay, employee involvement); business strategy such as increasing employees' skills, expanding into new markets) prevents us from examining if the effects of computers are felt differently by workers working for different types of employers. Also, a sizeable portion of the sample has missing information on annual income.

In the country as a whole, over 18% of workers who used computers stated that having to learn new computer skills caused them excess stress Everything else being equal, there does not appear to be any gender difference as male and female workers are equally likely to report stress caused by having to learn new computer skills (at 16.7%).

The likelihood for the need to learn new computer skills to cause stress in the workplace positively increases with age. The incidence of stress caused by having to learn new computer skills is estimated at 13% for workers under 35 and rises above 20% for workers over 45. This may in part be explained by the hypothesis that young workers are able to master computer skills faster/more easily than their older counterparts and hence feel less frustrated/stressed by the need to learn these skills.

Whether or not the worker has a university education makes a big difference. The estimated incidence of stress caused by having to learn new computer skills among workers with a university degree or beyond is only two-thirds of that estimated for workers whose education is below the university level (13.5% compared to 18.7%). This may also in part be explained by the hypothesis that better educated workers are able to master computer skills faster/more easily than their counterparts with lower education.

Foreign-born workers are more likely to report stress caused by having to learn new computer skills than those born in Canada (18.8% vs 16.3%). This may have something to do with the language barrier foreign-born workers face, especially among those newly arrived.

Stress caused by the need to learn new computer skills does not appear to be related to where a worker lives (an urban vs rural area or province), his/her work schedule (full-time relative to part-time), his/her employment type (paid work, self-employed with paid help, or own-account self-employed). But it does vary significantly with where a worker works in terms of industry and occupation. It is less likely to report this stress in accommodation and other services (around 14%), markedly more likely to experience it in education, management, finance and public administration (all over 18%), than in other industrial sectors (at 15%). By occupation, it is substantially more likely to report this stress in professional and processing occupations (above 20%) than in other professions (15%).

3.2 Work being affected

Overall, nearly 40% of workers who used computers reported that their work has been greatly affected by the introduction of computers, 21% somewhat affected, and the remaining 40% hardly or not at all affected. After controlling for other observables, male workers seem more likely to be affected by computers than their female counterparts. The

probability that work is greatly affected by computers is estimated at 40% for men compared to 33% for women. On the other hand, the likelihood that computers hardly or not at all affect work is around 35% for men compared to 41% for women.

The effect of computers on work appears to rise with age. The likelihood that work is greatly affected is estimated at 30% for the youngest group of workers, steadily rises for older groups and reaches 41% for those 45 to 54 years of age. On the contrary, the probability that computers hardly or not at all affect work is 45% for workers aged 15 to 24, gradually declines for older groups and reaches 34% those aged 45 to 54.

Computers have a significantly greater impact on better-educated workers. It is estimated that the work of 18% of workers with less than high school education is greatly affected by computers. This likelihood dramatically increases for better-educated workers and reaches 49% for those who have obtained at least a university degree. In contrast, the likelihood that computers hardly or not at all affect work is estimated at over 60% for workers with less than high school education, substantially drops for better-educated workers and reaches 26% for workers with at least a university degree.

Native-born workers are more likely than their foreign-born counterparts to be affected by computers. The likelihood that work is affected greatly is estimated at 38% and hardly or not at all at 36% for workers born in Canada. In comparison, the corresponding likelihood is 29% and 46% for workers born outside of the country, respectively. Computers do not affect work much differently across the provinces except for Alberta where a bigger impact is observed and for Newfoundland where a smaller impact is detected. Workers living in rural areas are slightly more likely to be affected by computers than their counterparts residing in urban areas. It is estimated that the work of 38% of workers living in rural areas is greatly affected by computers compared to that of 36% of those residing in urban areas. The exact opposite hold true with respect to hardly or no impact at all (36% for rural residents vs 38% for urban residents).

Work schedule makes a big difference — full-time workers are significantly more likely to be affected by computers than those working part-time. The probability that computers greatly affect work is estimated at 38% for full-time workers, 10 percentage points higher than that for their part-time counterparts. On the other hand, the likelihood that work is hardly or not at all affected by computers is estimated at 37% for those working full-time, 11 percentage points lower than that for part-time workers.

Computers have a smaller impact on the work of the own-account self-employed than that of the self-employed who hire others as well as that of wage and salary workers. On average, the probability that work is greatly affected by computers is 30% and hardly or not at all 45% for the self-employed who do not hire any paid help. In contrast, the corresponding likelihood is 37% and 37% for the self-employed with employees and regular paid employees, respectively.

The effect of computers varies significantly across industrial sectors. The most affected ones are finance and professional services where the work of over half of workers is affected greatly and under a quarter hardly or not at all. And the least impacted ones are construction, health and accommodation in which the work of under a quarter of workers is affected greatly and over half hardly or not at all.

There are also significant variations in the effects of computers on work by occupations. The estimated likelihood of work being greatly affected by computers ranges from a high of 51% in professional occupations to a low of under 20% in primary and processing professions. On the other hand, the probability that work is hardly or not at all affected is estimated at 25% for professional compared to around 60% for primary and processing professions.

All these results are not surprising as they, to a large degree, point to a positive association between the extent to which work is affected by computers and the extent/frequency of computer usage. When a characteristic is observed to be associated with a higher/more frequent use of computers, it is also identified to be associated with work being more affected; and vice versa (detailed analysis on incidence and frequency of computer use is provided in Lin and Popovic (2002a)).

Respondents who stated that their work has been greatly or somewhat affected by computers are further asked how their work is affected in terms of whether their job security has increased, decreased or stayed the same and whether their work has become more interesting, less interesting or stayed the same. Answers to these questions are also analyzed and what follows shows the results.

3.3 Has job become more/less secure?

Appendix A: (Continued)

Of those who stated that their work has been affected (greatly or somewhat) by the introduction of computers, 23% felt that their job security has increased, 68% thought that their job security has stayed the same, and the remaining 10% reported that their job security has decreased.

Everything else being equal, male workers benefit more from computers in terms of job security than their female counterparts. It is estimated that 22% of men are observed with a job security increase as a result of the introduction of computers compared to 16% among women. On the other hand, 8% of men are detected with a job security decrease relative to 12% for women.

While the impact of computers on job security is not correlated with workers' education attainment, it varies significantly across age groups and younger workers benefit more than their older counterparts. The probability that computers have increased job security is estimated at 28% for those aged 15 to 24, steadily declines for older groups and reaches less than half as high (13%) for the oldest group of workers. In comparison, the probability that job security has decreased is 6.2% for the youngest group of workers, gradually rises for older groups and reaches over twice as high (14%) for those aged 55 and over.

Foreign-born workers are affected by computers slightly more favorably in terms of job security change than their native-born counterparts. The likelihood that computers have made jobs more secure is 21% for the former, slightly higher than that of 19% for the latter, and the probability that computers have decreased job security is 8.7% for the former, slightly lower than that of 9.8% for the latter.

The impact of computers on job security does not differ much across the country except for two provinces. Compared to the rest of the country, the likelihood that job security has increased as a result of the introduction of computers is lower for Quebec (17% vs 20%) and the probability that jobs have become less secure higher (11.2% vs 9.5%). On the contrary, the probability that computers have made jobs more secure is higher for Alberta (at 24%) and the likelihood that job security has decreased lower (7.6%).

Whether a worker lives in a rural or an urban area is not associated with how his/her job security is affected by computers, nor is whether he/she is a regular wage and salary employee or self-employed with or without hiring any paid help. However, the number of hours he/she works on a weekly basis makes quite a difference. It is estimated that 20% of full-time workers felt that their jobs have become more secure as a result of the introduction of computers, 25% higher than that for part-time workers. On the other hand, 9.4% of those working full-time indicated that computers have made their jobs less secure, 25% lower than that for those working part-time.

The impact of computers on job security varies significantly across industries. Workers in manufacturing, agriculture, construction and accommodation gain the most as the probability that computers have made job more secure is the highest (nearly 25%) and that computers have made job less secure the lowest (at 7.3%). On the other hand, workers in finance and health services benefit the least as the estimated probability that computers have made job more/less secure is the lowest/highest (under 15% and over 13%, respectively).

The impact of computers on job security also differs markedly across occupations. The highest estimated probability that jobs have become more secure as a result of the introduction of computers is detected in the professional occupations (over 25%), over twice as high as in the trade professions (at 12%). The former is also observed with the lowest probability that computers have made jobs less secure (7%), under half of the highest also observed in the latter (over 15%).

These results demonstrate that while the magnitudes of computers' effects on job security differ from one group of workers to the other, substantially in some cases, the qualitative patterns observed above largely remain unchanged for both groups. That is, male workers benefit more than their female counterparts; younger workers profit more than older ones; workers in Quebec are disadvantaged while those in Alberta gain relative to the rest of the country. Full-time workers benefit more than those working part-time. In terms of industries and occupations, while the manufacturing, agriculture, construction and accommodation sectors benefit the most, the finance and health industries gain the least; the professional occupations gain the most, the trades professions benefit the least.

3.4 Has work become more/less interesting?

For the country as a whole, nearly six out of ten workers who stated that their work has been affected (greatly or somewhat) by the introduction of computers reported that their work has become more interesting as a result of the introduction of computers, over one-third reported that their work has become neither more nor less interesting, and 4% stated that their work has become less interesting. Controlling for other observable characteristics, women gain marginally more than men from computers in terms of work interest change. The likelihood that work has become more interesting as a result of the introduction of computers is 60% for women compared to 58% for men, and the probability that work has become less interesting is 3.6% for women compared to 4.0% for men.

Although education attainment does not make much of a difference, the impact of computers on work interest varies across age groups and younger workers gain more. The probability that computers have made work more interesting is estimated at 62% for those under 35 years of age, gradually declines to 56% for the oldest group. On the other hand, the likelihood that work has become less interesting as a result of the introduction of computers is 3.4% for those aged 15-34, gradually rises to 4.3% for those aged 55 and over.

Foreign-born workers benefit more from computers than their native-born counterparts. It is estimated that the work of 62% of the former has become more interesting as a result of the introduction of computers compared to 58% for the later. On the other hand, the likelihood that computers have made work less interesting is 3.3% for the former compared to 3.9% for the later.

The impact of computers on work interest does not differ with respect to where a worker lives, an urban or rural area or which province except for Alberta where the estimated probability that computers have made work more interesting is higher (66% vs 58%) and the probability that work has become less interesting lower (2.8% vs 4%) relative to other provinces.

The impact of computers on work interest does not vary whether a worker works full-time or part-time. Nor does it if he/she is a regular employee or a self-employed with paid help. However, those working on their own without hiring others benefit more. It is estimated that 63% of the own-account self-employed felt that their work has become more interesting as a result of the introduction of computers and 3.2% thought that their work has become less interesting compared to 58% and 3.9%, respectively, for employees and the self-employed employers.

There are significant industrial variations in the effects of computers on work interest. The estimated probability that computers have made work more interesting ranges from a low of 47% for health and 51% for construction and to a high of 63% for manufacturing, agriculture, forestry, professional services, management, information services, accommodation services, public administration and other services. And the contrary holds true for the likelihood that work has become less interesting as a result of the introduction of computers.

There are also significant variations in the impact of computers on work interest by profession. At the high end, the likelihood that work has become more interesting as a result of the introduction of computers is estimated at 64% for the managerial, professional and clerical occupations. At the other end of the scale, it is as low as 45% for processing and 49% for trades. The reverse is true for the probability that work has become less interesting.

To recap, those aged under 45 gain more from computers in terms of work interest change than their older counterparts; foreign-born workers are affected more favorably than their native-born counterparts; workers living in Alberta are advantaged and in British Columbia disadvantaged relative to the rest of the country; the own-account self-employed gain more than wage and salary workers as well as the self-employed employers. Breakdown by industry and occupation, the health services and transportation industries profit less relative to other sectors; and the trades, primary and processing professions benefit less relative to other occupations.

4. Summary and discussion

Computers have reached nearly every corner of our lives, whose impacts are inevitably wide-spread and profound. Does having to learn new computer skills cause extra stress in the workplace? The data on hand show that over 18% of computer-using workers thought so. Our regression results demonstrate that attributes that are significantly associated with workplace stress caused by the need to learn new computer skills include age, education, country of birth, industry and occupation. Specifically, having to learn new computer skills is more likely to cause workplace stress for older workers (e.g., workers aged 45 and over are nearly twice as likely to report this stress as those under 35). Workers with university education or beyond are less likely to experience this stress than their counterparts with below-university education. Foreign-born workers are more likely to report this stress than their native-born counterparts. It is less likely to report this stress in accommodation and other services, markedly more likely to experience it in education, management, finance and public administration. By occupation, it is substantially more likely to report this stress in professional and processing occupations.

Is work affected by the introduction of computers? The survey shows that 39% of workers reported that their work has been greatly affected, another 21% said that their work has been somewhat affected, while the remaining 40% felt that their work has been hardly or not at all affected. Our regression results reveal that characteristics that are significantly correlated to work being affected by computers include gender (greater impact on men), age (greater impact on older workers), education (greater impact on the better-educated), country of birth (greater impact on the native-born), area of residence (greater impact on those living in rural areas), work schedule (significantly greater impact on full-time workers), employment type (smaller impact on the own-account self-employed), industry (the most affected are finance and professional services and the least are health and accommodation), and occupation (professional occupations are the most affected and the primary and processing professions the least).

Has job become more or less secure as a result of the introduction of computers? Of those who stated that their work has been affected (greatly or somewhat), 23% felt that their job has become more secure, another 9% reported that their job has become less secure, while the majority (68%) thought that their job security has stayed the same. Our regression results demonstrate that observable attributes that are significantly correlated with job security change as a result of the introduction of computers include gender (men benefit more), age (younger workers benefit more), country of birth (foreign-born workers are affected more favourably), work schedule (full-time workers benefit more), industry (the manufacturing, agriculture, construction and accommodation sectors benefit the most, the finance and health industries gain the least), and occupation (the professional occupations gain the most, the trades professions benefit the least). This is largely as true for those who felt their work has been greatly affected by computers as for those who thought their work has only been somewhat affected.

Have computers made work more or less interesting? Of those who felt that their work has been affected, nearly six out of ten felt that their work has become more interesting, while 37% said that their work has become neither more nor less interesting and 4% stated that their work has become less interesting. Our regression results reveal that observable characteristics that are significantly associated with work interest change as a result of the introduction of computers include gender (women gain more), age (those aged under 35 benefit more), country of birth (foreign-born workers are affected more favourably), and employment type (the own-account self-employed benefit more). There are also significant variations across industry and occupation. By industry, health and construction benefit the least and manufacturing, agriculture, forestry, professional services, management, information services, accommodation services, public administration and other services are the bigger winners. Across occupations, the managerial, professional and clerical occupations benefit more and the processing and trades professions gain the least. These results apply, to a large extent, only to those who felt that their work has been greatly affected. There is not much variation across most of the explanatory variables for those who thought that their work has only been somewhat affected.

In short, our data clearly demonstrate that computers have profound impacts on the workplace — six out of ten workers feel that their work has been affected. Taken together, these results paint a pretty good-news picture of computer effects on job quality. Measured by job security (perceived by workers rather than reflected in actual statistics on turnover or job tenure/duration), winners outnumber losers by a ratio of 2.4 to 1. Measured by work interest, nearly fifteen workers feel that their work has become more interesting for every worker reporting that work has become less interesting. True, computers also have negative effects — nearly one out of five computer-using workers feel that having to learn new computer skills causes them extra stress at work.

There is also an issue of equity — not all workers are affected in the same way. There indeed exist substantial variations in these effects over demographic and job-specific characteristics. For example, older workers are affected more and they are affected less favourably. Some industries are affected more, and some industries are affected more favourably.

What do all of these imply then? While the effects of computers on job security and work interest may be the inherent nature of the computer revolution and there does not seem to be much individual workers, employers and public policy makers can do about them, there is certainly something we can do about the need to learn new computer skills as a workplace stressor.

Workplace stress can be caused by many factors: 1) factors related to the job (e.g., noise, boredom, shiftwork, fear to exposures to dangerous materials); 2) the role of the individual worker in the organization (e.g., insufficient information to perform tasks, lots

of responsibility but little authority and control); 3) social relationships and interpersonal demands; 4) prospects for promotion and advancement (e.g., inadequate recognition or reward for good performance); and 5) organizational structure and culture (e.g., inability to voice complaints or express feelings, prejudice) (Sutherland and Cooper (1988)). Now our results also show that the need to learn new computer skills is an important source of stress in the workplace. Among the few workplace stressors surveyed in the GSS, having to learn new computer skills constitutes the third biggest source of stress, far behind too many demands/hours and close to poor interpersonal relations.

Workplace stress can be very costly to both the employer and employee. In the short run, stress can lead to job dissatisfaction, which often results in absenteeism and reduced productivity. For example, Malik (1993) estimates that stress-related absenteeism costs the United States over $150 billion each year. Over the long run, stress can lead to health problems (e.g., heart disease, increased accident occurrence, poor mental health), substance abuse, and social/domestic problems (e.g., Friedman et al. (1996), Wheeler and Lyon (1992)).

Given all the negative outcomes of workplace stress, its reduction will be beneficial to workers, employers, and the society as a whole. As for having to learn new computer skills as a source of stress in the workplace, one effective way in reducing it is to equip workers with the required skills, be that general or specific.

Workers can explore various venues to learn and acquire these skills, be that formal or informal. Employers may encourage employees to do so by providing financial support and/or time off as well as providing direct training. From a public policy point of view, governments can encourage the population and employers to do so by providing financial incentives.

Training is generally regarded as an effective way for individuals to acquire various skills. We attempted in our modelling to address if participation in training helps reduce the incidence of workplace stress caused by the need to learn new computer skills. Unfortunately, we obtained no conclusive evidence largely due to the fact that nearly every computer-using worker took one form of training or another to learn computer skills. Future work (Lin, Carter and Popovic (2003)) will examine how Canadian workers acquire their computer skills, by way of formal training, on-the-job training, or through self-learning.

Appendix A: (Continued)

With all of these in mind, we close with a couple of caveats. First, we would like to stress that the effects of computers on job quality examined here are self-rated by survey respondents. There may very well be discrepancy between perceived effects and actual ones.

Second, the time period during which the survey is conducted (from January through December 2000) can be argued to be a very special phase of the business cycle. The overall economy, the high-tech sector in particular, has suffered a slowdown which resulted in massive layoffs ever since the completion of the survey. Coupled with accelerating advancements in computer and other advanced technologies, it may be reasonable to expect that responses could be different from what we have observed should the survey be conducted today. We therefore eagerly await data sources in the future to assess the impact of computers for different phases of the business cycle.

Appendix A: (Continued)

Context Summary Used in Experiment Two.

Information and communication technologies (ICTs),

nearly six out of ten Canadian workers used a computer (personal computer, mainframe or word processor) at work,

usage rate is up

one in two in 1993

earlier literature centers around the effects on productivity and job quality, which Rubery and Grimshaw (2001) sub-divide into three main dimensions: 1) employment relations and protection (e.g., employment opportunities, employment relations, career opportunities, job protection and collective bargaining, pay); 2) time and work autonomy (e.g., work intensity, power and autonomy, work/life balance, work relations); and 3) skills and careers (e.g., skills, job prospects).

the pessimistic argues that ICTs destroy employment opportunities through automation and rationalization, reduce pay by downgrading skills and weakening workers' collective bargaining power.

the optimistic hypothesizes that ICTs create jobs through developing new markets and human capital, increase pay by augmenting skills.

previous studies have found a positive relationship between productivity and the use of ICTs

there is a positive linkage between wages and the use of computers and other advanced technologies,

argued that workers who use computers earn more than other employees (those who do not) not because of their computing skills per se, but rather because they are endowed with more other unobservable or unmeasureable skills.

the adoption of computers and other new technologies is a key element to firms' success because these technologies are correlated with market share increases, productivity gains, product and delivery quality improvements, increased flexibility, production costs reduction, and so on

15% of unemployed job seekers in the United States used the internet to search for jobs in 1998,

Appendix A: (Continued)

internet job search rates exceeded those of traditional job search methods such

"The computer revolution has given the modern workplace an array of new options and improved efficiency. But far from having a calming effect on overworked employees, computerization has itself become a source of increasing psychological stress.

ICTs have increasingly replaced humans to perform a great number of complex and challenging tasks.

As a result, many processes and tasks have been automated or routinized

less effort and work on job quality as measured in psychological stress, job security and work interest

target population for this survey is all Canadians 15 years of age and older, who are not residents of the three territories (Yukon , Northwest and Nunavut) or full-time residents of institutions

GSS 2000 is a household-based survey

first question we try to address is whether having to learn new computer skills causes excess worry or stress in the workplace and

The second question we attempt to answer is whether work is affected by computers and if so, the extent to which work is affected.

Our third research question is thus how computers have changed job security

the fourth question addressed in the paper is how computers have changed work interest

the lack of information on employer characteristics (e.g.,

over 18% of workers who used computers stated that having to learn new computer skills caused them excess stress

the need to learn new computer skills to cause stress in the workplace positively increases with age

stress caused by having to learn new computer skills

young workers are able to master

Appendix A: (Continued)

stress caused by having to learn new computer skills among workers with a university degree or beyond is only two-thirds of that estimated for workers whose education is below the university level

hypothesis that better educated workers are able to master computer skills faster/more easily than their counterparts with lower education.

Foreign-born workers are more likely to report stress caused by having to learn new computer skills than those born in Canada (18.8%

Stress caused by the need to learn new computer skills does not appear to be related to where a worker lives (an urban vs rural area or province), his/her work schedule (full-time relative to part-time), his/her employment type (paid work, self-employed with paid help, or own-account self-employed).

But it does vary significantly with where a worker works in terms of industry and occupation

By occupation, it is substantially more likely to report this stress in professional and processing occupations (above 20%) than in other professions (15%).

nearly 40% of workers who used computers

21% somewhat affected

the remaining 40% hardly or not at all affected.

The effect of computers on work appears to rise with age.

Computers have a significantly greater impact on better-educated workers

Native-born workers are more likely than their foreign-born counterparts to be affected by computers.

Workers living in rural areas are slightly more likely to be affected by computers than their counterparts residing in urban areas.

Work schedule makes a big difference □ full-time workers are significantly more likely to be affected by computers than those working part-time.

Computers have a smaller impact on the work of the own-account self-employed than that of the self-employed who hire others as well as that of wage and salary workers.

Appendix A: (Continued)

effect of computers varies significantly across industrial sectors. The most affected ones are finance and professional services where the work of over half of workers is affected greatly and under a quarter hardly or not at all.

the least impacted ones are construction, health and accommodation in which the work of under a quarter of workers is affected greatly and over half hardly or not at all.

work by occupations

positive association between the extent to which work is affected by computers and the extent/frequency of computer usage.

work has been affected (greatly or somewhat) by the introduction of computers

male workers benefit more from computers in terms of job security than their female counterparts.

impact of computers on job security is not correlated with workers' education attainment, it varies significantly across age groups and younger workers benefit more than their older counterparts.

Foreign-born workers are affected by computers slightly more favorably in terms of job security change than their native-born counterparts.

impact of computers on job security does not differ much across the country except for two provinces.

lower for Quebec

higher for Alberta

Whether a worker lives in a rural or an urban area is not associated with how his/her job security is affected by computers

the number of hours he/she works on a weekly basis makes quite a difference

more secure

impact of computers on job security varies significantly across industries

manufacturing, agriculture, construction and accommodation gain the most

differs markedly across occupations

Appendix A: (Continued)

highest estimated probability that jobs have become more secure as a result of the introduction of computers

male workers benefit

workers in Quebec are

disadvantaged while

Alberta gain relative to the rest of the country.

manufacturing, agriculture, construction and accommodation sectors benefit the most

finance and health industries gain the least

professional occupations gain the most,

trades professions benefit the least.

nearly six out of ten workers who stated that their work has been affected (greatly or somewhat) by the introduction of computers reported that their work has become more interesting as a result of the introduction of computers, over one-third reported that their work has become neither more nor less interesting, and 4% stated that their work has become less interesting.

women gain marginally more than men from computers in terms of work interest change.

impact of computers on work interest varies across age groups

Foreign-born workers benefit more from computers than their native-born counterparts.

work interest does not differ with respect to where a worker lives

working on their own without hiring

significant industrial variations

significant variations in the impact of computers on work interest by profession

those aged under 45 gain more from computers in terms of work interest change than their older counterparts;

foreign-born workers are affected more favorably than their native-born counterparts;

Appendix A: (Continued)

workers living in Alberta

British Columbia disadvantaged

own-account self-employed

attributes that are significantly associated with workplace stress caused by the need to learn new computer skills include age, education, country of birth, industry and occupation.

learn new computer skills

more likely to cause workplace stress for older workers

university education or

Foreign-born workers are more likely to report this stress

Is work affected by the introduction of computers? The survey

39% of workers reported that their work has been greatly affected, another 21% said that their work has been somewhat affected, while the remaining 40% felt that their work has been hardly or not at all affected.

attributes that are significantly correlated with job security change

work has been affected,

nearly six out of ten felt that their work has become more interesting

data clearly demonstrate that computers have profound impacts on the workplace

six out of ten workers feel that their work has been affected.

Workplace stress can be caused by many factors: 1) factors related to the job (e.g., noise, boredom, shiftwork, fear to exposures to dangerous materials); 2) the role of the individual worker in the organization (e.g., insufficient information to perform tasks, lots of responsibility but little authority and control); 3) social relationships and interpersonal demands; 4) prospects for promotion and advancement (e.g., inadequate recognition or reward for good performance); and 5) organizational structure and culture (e.g., inability to voice complaints or express feelings, prejudice) stress can lead to job dissatisfaction, which often results in absenteeism and reduced productivity.

Training is generally regarded as an effective way for individuals to acquire various skills.

Appendix A: (Continued)

Keyword Summary used in Experiment Two.

Information and communication technologies (ICTs),

nearly six out of ten Canadian workers used a computer

usage

up

effects on productivity and job quality,

1) employment relations and protection

time and work autonomy

3) skills and careers

ICTs destroy employment opportunities through automation and rationalization,

downgrading

weakening

optimistic

create

developing new markets

human capital,

positive relationship

productivity and the use of ICTs

positive association between wages and the use of ICTs

positive

workers who use computers earn more than other employees

endowed

more

unobservable or unmeasureable skills.

adoption

computers

new technologies is a key element to firms' success

technologies

market share increases,

productivity gains,

product and delivery quality improvements,

increased flexibility,

production costs reduction,

15%

unemployed job seekers in the United States used the internet to search for jobs in 1998,

internet job search rates exceeded those of traditional job search methods

options

improved efficiency.

computerization has itself become a source of increasing psychological stress."

ICTs

replaced humans

Canadians 15 years

older,

GSS 2000

first question

having to learn new computer skills causes excess worry or stress in the workplace

varies

observable demographic attributes,

locations,

work characteristics.

whether work is affected by computers

extent

affected.

how computers have changed job security.

increased";

same";

decreased".

how computers have changed work interest.

limitations.

lack

information

employer characteristics

18% of workers who used computers stated that having to learn new computer skills caused them excess stress

gender difference

likelihood for the need to learn new computer skills to cause stress in the workplace positively increases with age.

Whether or not the worker has a university education makes a big difference

stress

degree

educated

skills

Foreign-born workers are more likely to report stress

computer skills

language barrier

Stress caused by the need to learn new computer skills does not appear to be related to where a worker lives

schedule

employment type

vary

where

worker works in terms of industry and occupation.

finance and

professional and processing occupations

40%

workers

computers

work has been greatly affected

computers,

21% somewhat affected,

40% hardly or not at all affected.

male workers seem more likely to be affected by computers than their female counterparts.

effect of computers on work appears to rise with agee

affected

workers,

Computers have a significantly greater impact on better-educated workers.

affected

Native-born workers are more likely than their foreign-born counterparts to be affected by computers.

Workers

rural areas are slightly more likely

affected

computers

Work schedule makes a big difference

full-time workers are significantly more likely to be affected by computers than those working part-time.

Computers have a smaller impact on the work of the own-account self-employed than that of the self-employed who hire others as well as that of wage and salary workers.

varies significantly across industrial sectors.

most affected ones are finance and professional services

least impacted

construction, health and accommodation

professional

processing

positive association

extent

work is affected

computers

extent/frequency of computer usage.

23% felt that their job security has increased,

security

same,

security

decreased.

male workers benefit more from computers in terms of job security than their female counterparts.

security

education

younger workers benefit more than their older counterparts.

Foreign-born workers

affected

slightly more

job security

Quebec

Appendix A: (Continued)

Alberta

rural

urban

not associated

job security

hours

quite a difference.

job security varies

significantly

industries.

Workers in manufacturing, agriculture, construction and accommodation gain the most as

probability

computers have made job more secure

workers in finance and health services benefit the least

more/less secure

job security

occupationss

is detected in the professional occupations (over

trade

male workers benefit more than their female counterparts; younger workers profit more than older ones; workers in Quebec are disadvantaged while those in Alberta gain relative to the rest of the country. Full-time workers benefit more than those working part-time. In terms of industries and occupations, while the manufacturing, agriculture, construction

Appendix A: (Continued)

and accommodation sectors benefit the most, the finance and health industries gain the least; the professional occupations gain the most, the trades professions benefit the least.

six out of ten workers

more interesting

women

AAlthougheducation attainment

not make

differencee

more.

Foreign-born workers benefit more from

native-born

work interest does

not differ

where a

worker lives,

impact

computers

work interest

not vary

works full-time or part-time.

working on their own

benefit more.

significant industrial variations

interest.

low of 47% for health and 51% for construction and

high of 63% for manufacturing, agriculture, forestry, professional services, management, information services, accommodation services, public administration and other services.

professionn

managerial,

professional

clerical

processing

trades.

those aged under 45 gain more from computers in terms of work interest change than their older counterparts; foreign-born workers are affected more favorably than their native-born counterparts;

workers living in Alberta are advantaged and in British Columbia disadvantaged relative to the rest of the country;

the own-account self-employed gain more than wage and salary workers as well as the self-employed employers.

health services and transportation industries profit less

18%

workers thought so.

stress

learn

new

Appendix A: (Continued)

skills

more

workplace stress for older

university education

less likely

stress

Foreign-born

more

professional

processing occupations.

39%

greatly affected,

40%

hardly

not

affected.

correlated

work being affected

gender (greater impact on men),

age (greater impact on older workers),

education (greater impact on the better-educated),

country of birth (greater impact on the native-born), area of residence (greater impact on those living in rural areas),

work schedule (significantly greater impact on full-time workers),

employment type (smaller impact on the own-account self-employed),

industry (the most affected are finance and professional services and the least are health and accommodation),

occupation (professional occupations are the most affected and the primary and processing professions the least).

majority (68%) thought that their job security has stayed the same.

gender

age

country of birth

schedule

industry

occupation

six out of ten

more interesting,

health

construction

least

more

impacts

workplace

six out of ten

affected.

learn

skills

stress

Workplace stress can be caused by many factors:

factors related to the job

role

individual worker

social relationships and interpersonal demands;

prospects

promotion

advancement

organizational structure and culture

new computer skills

third biggest source

stress,

equip workers

Training is

effective

various skills

Appendix B: Experiment One Registration Screen

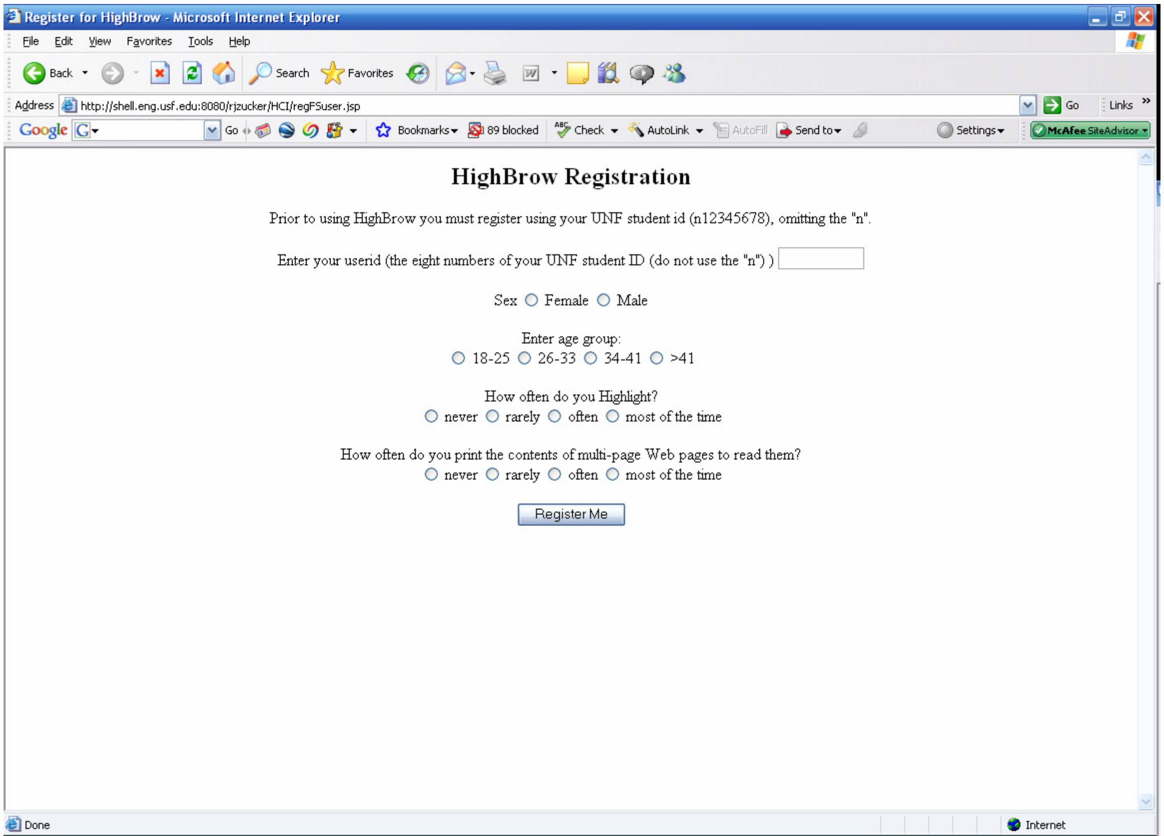


Figure B.1 - Experiment One Registration Screen

Appendix C: HighBrow Installation Process Instructions

Directions for installing HighBrow 2.0, the context highlighting browser for Intro to OOP.

Quick Install:

Building HighBrow2.0 (approx time 1 minute)

The next step will help us install Highbrow 2.0.

Save the HighBrow2 zip file onto your machine (save it to your Desktop), then extract the files to a directory called C:\HB. Directions on how to extract from a zip file.

In the C:\HB directory, double click on the **buildit** file, this will copy underlying software to appropriate directories and launch HighBrow 2.0. This need only be done once.

If you can successfully get the Computer Applications for Business HighBrow webpage then the installation is complete and to run it from here on out all you need do is double click on the runit file in the c:\HB directory. If it doesn't work, then the Java Runtime Environment must be installed. Please follow the **If Quick Install did not work** steps below (approx time 5-10 minutes):

Creating a Shortcut (optional, approx time 1 minute)

You may create a shortcut by right clicking the runit file and selecting Create Shortcut. This Shortcut may be placed on the desktop by dragging the shortcut to the desktop.

If Quick Install did not work:

Verifying the JRE:

Highbrow 2.0 runs on the java 1.5 platform, and is not compatible with lower versions of java. In order to determine which version of java you are running, open the Windows Explorer (click the windows button and the E simultaneously). Locate and open the directory called Program Files (found under the C: directory). Now locate Java, if Java is not found you may proceed to the **Installing the current JRE** step, otherwise click on Java (for a pictorial please see picture of directory tree). You should see a jre.1.5.0_06 or better (better means the 1.5.0_06 have higher values). If your jre (Java Runtime Environment) is less than jre 1.5, then please remove it (next step) otherwise skip to building HighBrow2.0.

Removing an older version of the JRE:
Please go to control panel, then "Add or remove programs" and remove the existing J2SE runtime environment. Now you must install the current version of the jre.

Installing the current JRE:
This step may take a while (especially if you have a dialup access). We must now install the jre from the java.sun.com web site (http://java.sun.com/j2se/1.5.0/download.jsp) or simply clicking on this link: **Download JRE 5.0 Update 6**. Accept the license, then select "Windows Offline Installation, Multi-language" (the 16.0 MB version) ***VERY IMPORTANT*****:** Please install the jre in the default directory specified by the installer. This is important as the HighBrow **buildit** command will use that directory for installing a key component.

If you are a developer of Java programs, then you may want to download the Java Development Kit (jdk). There is no need to download both as the jdk includes the jre as well as other programs used in the development process.

Building HighBrow2.0:
The next step will help us install Highbrow 2.0 In the C:\HB directory, double click on the buildit file (it is important to run buildit, this second time, after you have installed the new JRE), this will copy underlying software to appropriate directories and launch HighBrow 2.0. This need only be done once. If you are still not successful please contact me by phone 636-5935 or via e-mail (rzucker@unf.edu) or via Blackboard.

Running HighBrow2.0:
After running buildit, the only thing you have to do is click on the runit in the C:\HB directory.

Creating a Shortcut (optional)
You may create a shortcut by right clicking the runit file and selecting Create Shortcut. This Shortcut may be placed on the desktop by dragging the shortcut to the desktop.

Good luck and enjoy HighBrow 2.0.

Appendix D: Test Used for Experiment One and Two

Extra Credit Quiz

Userid:_________________________

Instructions: The following quiz contains twenty multiple choice questions. Please **circle** the best answer based on the survey results from your reading. Each question is worth 1.5 points. This quiz is closed book and closed notes.

1) The “Effects of Computers” study was performed in
   a) Canada
   b) Mexico
   c) U.S.A.

2) Learning new computer skills is the ____________ biggest source of stress in the workplace.
   a) First.
   b) Second.
   c) Third.
   d) Fourth.

3) Using age as the sole criteria, which age group is more likely to experience workplace stress caused by learning new computer skills.
   a) Less than 20 years old.
   b) 25 to 32 years old.
   c) 34 to 40 years old.
   d) Greater than 45 years old.

4) Which of the following is **not** an attribute that is significantly associated with workplace stress caused by the need to learn new computer skills?
   a) Industry.
   b) Education.
   c) Income.
   d) Country of birth.
   e) All of the above attributes are associated with workplace stress.

5) Foreign-born workers are ________________ to report stress caused by having to learn new computer skills than native-born.
   a) Less likely.
   b) Equally likely.
   c) More likely.
   d) Occupation was not considered an attribute in reporting stress.

6) With respect to the likelihood persons in different occupations reporting workplace stress, which of the following statements is true?
   a) There is no difference in the likelihood to report stress across occupations.
   b) Accommodations and other services are more likely to report stress than professional and processing occupations.
   c) Accommodation and other services are less likely to report stress than professional and processing occupations.
   d) Accommodation and professional are less likely to report stress than processing and other services occupations.
   e) Occupation was not considered an attribute in reporting stress.

7) Which respect to gender as a factor in reporting workplace stress, which of the following statements is true?
   a) There is no difference in the likelihood to report stress across genders.
   b) Males are more likely to report stress than females.
   c) Females are more likely to report stress than males.
   d) Gender was not considered an attribute in reporting stress.

8) With respect to impact of work being affected by computers, which is not included in the study?
   a) Industry
   b) Education
   c) Gender
   d) Country of birth
   e) All of the above items were included in the study.

9) With respect to location and impact of work being affected by computers:
   a) Computers do not affect work much differently across the provinces except for Alberta and Newfoundland.
   b) Computers affect work differently across provinces, with Quebec having the greatest affect and British Columbia having the least.
   c) Computers affect work differently across provinces with the western provinces having the greatest affect in the eastern provinces having the least.
   d) Computers affect work differently across provinces with no discernible pattern.
   e) Location and the affect of computers on work were not covered in the study.

10) With respect to the impact of computers on work and the area of residence:
   a) Persons living in urban areas had greatest impact.
   b) Persons living in suburban areas had greatest impact.
   c) Persons living in rural areas had greatest impact.
   d) The area of residence and the impact of computers on work was not covered in the study.

Appendix D: (Continued)

11) With respect to the impact of computers on work and employment type:
   a) Computers have a higher impact on the work of the own-account self-employed than that of self-employed who hire others.
   b) Computers have a higher impact on the work of the self-employed who hire others then that of the own-account self-employed.
   c) Computers offered no significant difference in impact on the work of the self employed (own-account or those who hire others)
   d) Employment type and the impact of computers on work was not covered in the study.

12) The association to which work is affected by computers and the extent/frequency of computer usage is:
   a) Directly proportional.
   b) Inversely proportional.
   c) No pattern exists in relation.

13) The survey reported that the impact of computers on job security was
   a) Over twice as high for professional occupations as the trade professions.
   b) Over twice as high for trade professions as the professional occupations.
   c) No significant differences reported across occupations

14) Overall the group of workers benefiting most in terms of job security are:
   a) Younger, females
   b) Younger, males
   c) Older, males
   d) Older, females.

15) The estimated probability that computers have made work more interesting was lowest in the ________ industry.
   a) agriculture
   b) construction
   c) health
   d) manufacturing

16) Computers have a significantly greater impact on the work of
   a) workers with less than a high school education
   b) workers with a minimum high school education
   c) workers who have obtained a university degree
   d) All levels of education are equally affected.

Appendix D: (Continued)

17) The majority of workers who stated that their work has been affected by the introduction of computers reported that their work has become
    a) less interesting as a result of the introduction of computers.
    b) more interesting as a result of the introduction of computers.
    c) neither more nor less interesting.

18) The gender that shows the most interest gain as a result of the introduction of computers.
    a) Males.
    b) Females.
    c) Neither gender gained interest.

19) The occupation that gains the least interest from the introduction of computers is:
    a) Clerical.
    b) Managerial.
    c) The trades profession.
    d) The sports profession.

20) The paper did offer a way to reduce the stress involved with learning new computer skills in the workplace.
    a) True.
    b) False.

Thank you for your participation!

Appendix E: Human Research Informed Consent Form

University of North Florida
College of Computing, Engineering, & Construction
Computer and Information Sciences Department

University of South Florida
Social and Behavioral Sciences

Human Research Informed Consent Form

1. **Title of Study:** HighBrow: A context enabled web browser.
2. **Principal Investigator:** Ron Zucker.
3. **Study Location:** University of North Florida.
4. **Statement of participation:** Participation in this study is voluntary; refusal to participate will involve no penalty or loss of benefits to which the participant is otherwise entitled. The data will be collated and summarized; there is no intention to study individual performance. The participants may discontinue participation at any time without penalty or loss of benefits to which the participant is otherwise entitled.
5. **Statement of confidentiality:** Your privacy and research records will be kept confidential to the extent of the law. Authorized research personnel, employees of the Department of Health and Human Services, and the USF Institutional Review Board and its staff, and any other individuals acting on behalf of USF, may inspect the records from this research project.

   The results of this study may be published. However, the data obtained from you will be combined with data from others in the publication. The published results will not include your name or any other information that would personally identify you in any way.
6. **Questions relating to the study:** Should you have questions concerning this study please contact any or all of the following:
   a. Principal investigator: Ron Zucker, 15/3227, (904) 620-1329, rzucker@unf.edu
   b. Adviser: Dr. Rundus, USF, (813) 974-4184, rundus@csc.usf.edu
   c. Institutional review board: Dr. Kathaleen Bloom, Chair, UNF Institutional Review Board, (904) 620-2084
   d. Division of Research Integrity and Compliance of the University of South Florida at (813) 974-5638.
7. **Brief description of study:** this study involves research in the area of human computer interaction using a software prototype of a context highlighting browser. The purpose of this study is to determine the effectiveness of context highlighting for the originator of the highlighting and for consumers of previously highlighted material. The study will take place in a period not to exceed seven weeks. The participant will read and highlight selected articles, take a quiz on the articles, and fill out a brief questionnaire concerning the use of HighBrow. Anticipated time required for participation is under three hours.
8. **Description of risk:** This is a minimal risk study. Any risk encountered would be equivalent to that of normal computer use.

9. **Compensation for participation:** Participation in the study is purely voluntary and is not compensated. Extra credit will be given to students participating in this experiment. The points will be normalized between the control (keyword only) and the context (keyword and context) groups to insure fairness.

   Students who do not wish to participate in the study will be allowed to receive the equivalent extra credit points by writing a paper on the topics covered in the study
10. **Minimum age**: Participants in the study must be at least eighteen years of age.

You may get further information about UNF policies, the conduct of the study, the rights of research subjects or if you suffer injuries related to your participation in this research project, from the Chair of the Institutional Review Board, Dr. Kathaleen Bloom at (904) 620-2684. If you have any questions about your rights as a person taking part in a research study, you may contact the Division of Research Integrity and Compliance of the University of South Florida at (813) 974-5638.

I have had an opportunity to have my questions answered. I have been given a copy of this form. I agree to take part in the study. I'm over in eighteen years of age.

I am at least eighteen years old. ______ (initial)

I have had the study that I am agreeing to participate in explained to me to my satisfaction. ______ (initial)

I have had the opportunity to ask any questions I may have had regarding the study. ______ (initial)

I agreed to participate in the HighBrow: the context sensitive highlighting browser study being conducted by Mr. Ronald J. Zucker, the University of North Florida and the University of South Florida.

____
Participant Printed Name

____ ____________
Participant Signature Date

__Ronald J. Zucker______________________________
Principal Investigator Printed Name

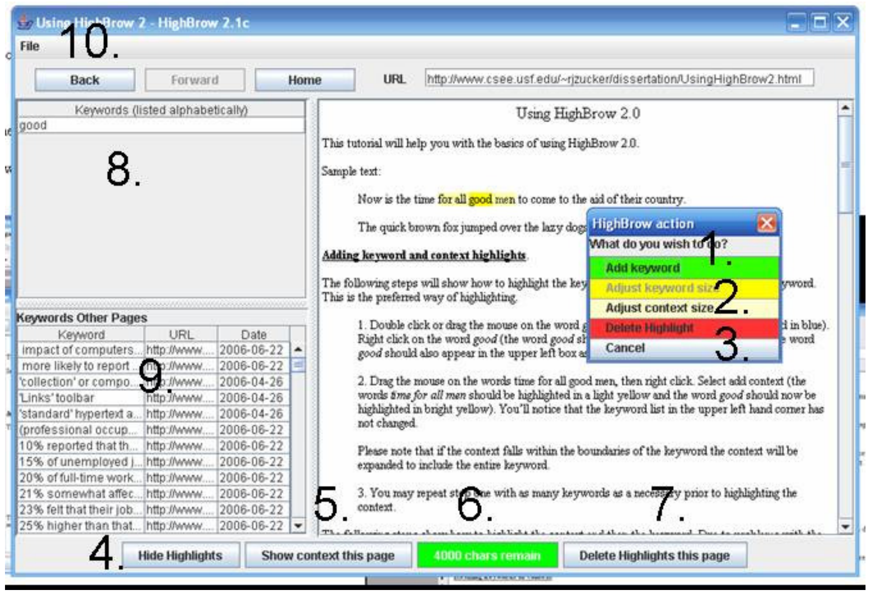
____ ____________
Principal Investigator Date

Original copy-Principal investigator/duplicate copy-participant

Appendix F: HighBrow Instructions (Context/Keyword)

Using HighBrow 2.0

This tutorial will help you with the basics of using HighBrow 2.0. The following figure is a screen shot of HighBrow (the numbers correspond to the items listed below).



Sample text:

> Now is the time for all good men to come to the aid of their country.
>
> The quick brown fox jumped over the lazy dogs.

<u>1. Adding keyword and context highlights</u>.

The following steps will show how to highlight the keyword and the context surrounding the keyword. This is the preferred way of highlighting.

> A. Using the sample text above, double click or drag the mouse on the word *good* (the word *good* should be highlighted in blue). Right click on the word *good* (the word *good* should be highlighted in bright yellow). The word *good* should also appear in the upper left box as a keyword. Note Highbrow will assume that this is a keyword selection and automatically highlight in bright yellow.
>
> B. Again using the sample text above, drag the mouse on the words time for all good men, then right click. Select add context from the pop-up menu (the words *time for all men* should be highlighted in a light yellow and the word *good* should

now be highlighted in bright yellow). You'll notice that the keyword list in the upper left hand corner has not changed.

Please note that if the context falls within the boundaries of the keyword the context will be expanded to include the entire keyword.

C. You may repeat step one with as many keywords as a necessary prior to highlighting the context.

The following steps show how to highlight the context and then the keyword. Due to problems with the interface this method is not recommended because no feedback is given when highlighting a keyword.

A. Using the sample text above, drag the mouse across *quick brown fox jumped over the lazy*, then right click (this should now show *quick brown fox jumped over the lazy* as a keyword highlighted in bright yellow and also appearing in the upper left hand box as a keyword).

B. Now in the sample text above, drag the mouse over the word *fox* (you should notice that no highlighting takes place). Right click and select add context. The word *fox* should now be highlighted in bright yellow indicating a keyword and the words *quick brown jumped over the lazy* are now in light yellow, indicating context. You should also notice that in the upper left hand box, the phrase *quick brown fox jumped over the lazy* has been replaced with the keyword *fox*.

C. You may add as many keywords as necessary in the context by repeating step 2.

2, Resizing keywords or context.

To resize a keyword, select the new beginning and ending point then right click and select resize keyword. **Please note:** The keyword will be sized to the new beginning and the new end. If you wish to add to the existing highlight you must start at either the existing beginning or end at the existing end. If the resizing encompasses other keywords these keywords may be swallowed by the resized keyword (i.e., the words swallowed will be deleted as separate keywords).

Resizing context works in a similar manner to resizing keywords. The only difference is that the context will be lengthened to include the entire keyword if a keyword either begins or ends the context. If the context becomes smaller than the keyword the context is effectively deleted.

3. Deleting keywords or context.

In order to delete a keyword, simply select the entire keyword or any portion thereof, right click then select delete. **Please note:** If you delete the last keyword within a context the context will become the keyword.

To delete context, select a portion of the context outside of a keyword then right click and select delete. Make sure that no keywords are selected in this process.

4, Showing and hiding highlights.

The button on the bottom left will allow you to either hide or show highlights on this page. This button will only be activated when there are highlights. This button will not delete any highlights and is used for display purposes only.

5. Show context this page.

The show context this page button located to the right of the show and hide highlights button is used to produce the context summary (see item 11. below) including highlighted keywords for the current page. Please note that altering highlights will not automatically update the context page, however, reclicking on the show context this page button will refresh the summary.

6. Highlight status

The third button from the left on the bottom is a status of the amount of key characters they can be highlighted in any one highlight. This is due to limitations in the database. You may highlight, though not recommended, several 4000 character sections (remember the idea here is to summarize, not copy, the original document). If you choose to select more than 4000 characters, you'll be issued a warning and the data will not be stored.

7. Deleting all highlights.

The button on the lower right indicates delete highlights this page. This button will permanently delete all highlights (keyword and context) on this page. Restoring the highlights will have to be done manually so please use caution when using this button.

8. KeyWords (listed alphabetically)

The box located to the left and above the Keywords Other Pages box lists the keyword for the current page in alphabetical order. Selecting a keyword row in this table will bring the keyword (and the resulting context) into view. The keyword box may be resized by dragging the right or bottom borders.

9. Keywords Other Pages

The box located on the left and below the keyword box represents keywords found on other pages. This data is sorted on the keyword (alphabetically and is the default). You may change the sort order by clicking on the titles (URL (ascending) or Date (descending)). This box may be resized by dragging the right or top borders. The columns of the table may also be resized by adjusting the titles in a similar manner. Clicking on a row (other than the title row) will produce the context summary (see item 11. below) for that row. To see the original webpage, you may select “Go to Webpage” on the menu bar of the context summary.

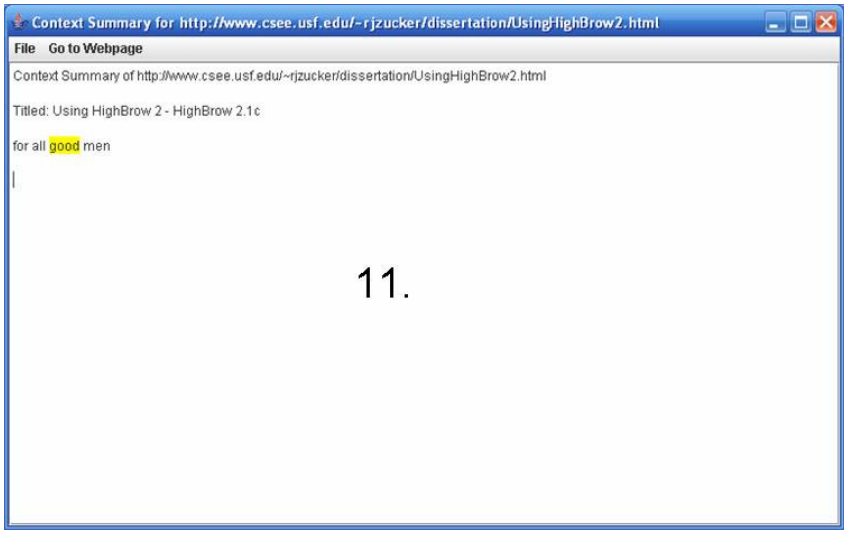
10. Printing the Highlighted Document

The current document may be printed by selecting the File menu and selecting the Print button. The document will be printed showing the highlights (provided they are not hidden).

11. Context Summary

The context summary page (obtained by clicking on the “Show context this page” button (item 5. above) selecting a keyword from the keywords other pages (item 9. above))shows the context and highlighted keywords (note context is not highlighted as it would be considered redundant). The summary is presented in the same order as it appears in the original document. The summary may be printed or copied to a file by selecting the File menu and clicking on either Print Context Summary or Copy Context. The copied summary will not include the highlights (but will include the highlighted words), the printed copy will print the highlights. Exiting or Closing the Context Summary will not shutdown HighBrow it will simply close the Context Summary. To avoid confusion, only one context summary is visible at a time.

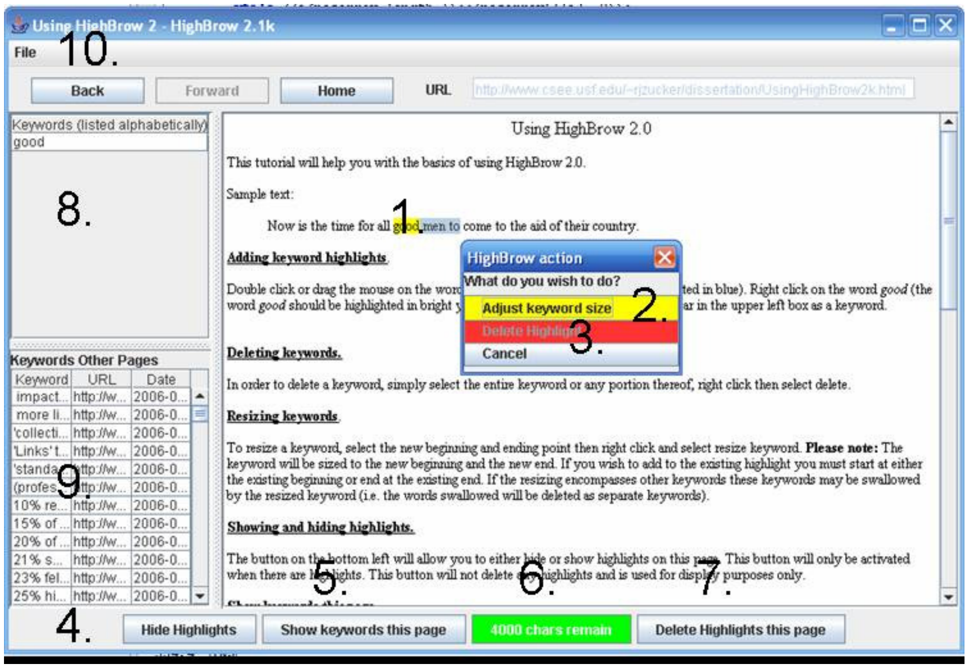
Appendix F: (Continued)



Appendix G: HighBrow Instructions (Keyword Only)

Using HighBrow 2.0

This tutorial will help you with the basics of using HighBrow 2.0. The following figure is a screen shot of HighBrow (the numbers correspond to the items listed below).



Sample text:

> Now is the time for all good men to come to the aid of their country.

1. Adding keyword highlights.

Using the Sample text above, double click or drag the mouse on the word *good* (the word *good* should be highlighted in blue). Right click on the word *good* (the word *good* should be highlighted in bright yellow). The word *good* should also appear in the upper left box as a keyword.

2. Resizing keywords.

To resize a keyword, select the new beginning and ending point then right click and select resize keyword. **Please note:** The keyword will be sized to the new beginning and the new end. If you wish to add to the existing highlight you must start at either the existing beginning or end at the existing end. If the resizing encompasses other keywords these keywords may be swallowed by the resized keyword (i.e., the words swallowed will be deleted as separate keywords).

3. Deleting keywords.

In order to delete a keyword, simply select the entire keyword or any portion thereof, right click then select delete.

4. Showing and hiding highlights.

The button on the bottom left will allow you to either hide or show highlights on this page. This button will only be activated when there are highlights. This button will not delete any highlights and is used for display purposes only.

5. Show keywords this page.

The show keywords this page button located to the right of the show and hide highlights button is used to produce a summary of highlighted keywords (See item 11. below) Please note that altering highlights will not automatically update the summary page, however, reclicking on the show keywords this page button will refresh the summary.

6. Highlight status

The third button from the left on the bottom is a status of the amount of key characters they can be highlighted in any one highlight. This is due to limitations in the database. You may highlight, though not recommended, several 4000 character sections (remember the idea here is to summarize, not copy, the original document). If you choose to select more than 4000 characters, you'll be issued a warning and the data will not be stored.

7. Deleting all highlights.

The button on the lower right indicates delete highlights this page. This button will permanently delete all highlights (keyword and context) on this page. Restoring the highlights will have to be done manually so **please use caution when using this button**.

8. KeyWords (listed alphabetically)

The box located to the left and above the Keywords Other Pages box lists the keyword for the current page in alphabetical order. Selecting a keyword row in this table will bring the keyword (and the resulting context) into view.

9. Keywords Other Pages

The box located on the left and below the keyword box represents keywords found on other pages. This data is sorted on the keyword (alphabetically and is the default). You may change the sort order by clicking on the titles (URL (ascending) or Date (descending)). The frames may be resized by dragging the boundaries in the direction

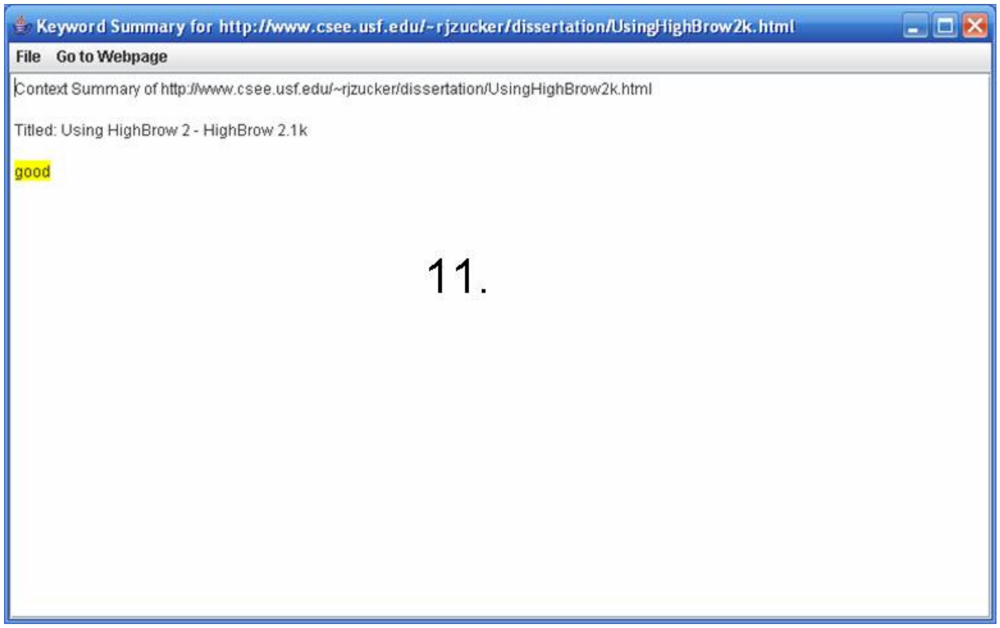
you wish to size. The columns of the table may also be resized by adjusting the titles in a similar manner. Clicking on a row (other than the title row) will produce the keyword summary (see item 11. below) for that row. To see the original webpage, you may select "Go to Webpage" on the menu bar of the keyword summary.

10. Printing the Highlighted Document

The current document may be printed by selecting the File menu and selecting the Print button. The document will be printed showing the highlights (provided they are not hidden).

11. Keword Summary

The keyword summary page shows the highlighted keywords (see items 5 and 9). The summary data is presented in the same order as it appears in the original document. The summary may be printed or copied to a file by selecting the File menu and clicking on either Print Context Summary or Copy Context. The copied summary will not include the highlights (but will include the highlighted words), the printed copy will print the highlights. Exiting or Closing the Context Summary will not shutdown HighBrow it will simply close the Context Summary. To avoid confusion, only one context summary is visible at a time.


Keyword Summary for http://www.csee.usf.edu/~rjzucker/dissertation/UsingHighBrow2k.html
File Go to Webpage
Context Summary of http://www.csee.usf.edu/~rjzucker/dissertation/UsingHighBrow2k.html
Titled: Using HighBrow 2 - HighBrow 2.1k
good
11.

Appendix H: Test Score Tabulation

Experiment One: Context/Keyword (Times Not Taken)

Group	Score
Context/Keyword Experiment One	10
Context/Keyword Experiment One	17
Context/Keyword Experiment One	16
Context/Keyword Experiment One	8
Context/Keyword Experiment One	7
Context/Keyword Experiment One	12
Context/Keyword Experiment One	11
Context/Keyword Experiment One	12
Context/Keyword Experiment One	15
Context/Keyword Experiment One	10
Context/Keyword Experiment One	12
Context/Keyword Experiment One	12
Context/Keyword Experiment One	12
Context/Keyword Experiment One	14
Context/Keyword Experiment One	16
Context/Keyword Experiment One	13
Context/Keyword Experiment One	18
Context/Keyword Experiment One	9
Context/Keyword Experiment One	12

Experiment One: Keyword Only (Times Not Taken)

Group	Score
Keyword Only Experiment One	11
Keyword Only Experiment One	10
Keyword Only Experiment One	14
Keyword Only Experiment One	6
Keyword Only Experiment One	12
Keyword Only Experiment One	14
Keyword Only Experiment One	11
Keyword Only Experiment One	11
Keyword Only Experiment One	8
Keyword Only Experiment One	11
Keyword Only Experiment One	9
Keyword Only Experiment One	9
Keyword Only Experiment One	12
Keyword Only Experiment One	13
Keyword Only Experiment One	7
Keyword Only Experiment One	14
Keyword Only Experiment One	15
Keyword Only Experiment One	12
Keyword Only Experiment One	13
Keyword Only Experiment One	9
Keyword Only Experiment One	14
Keyword Only Experiment One	13
Keyword Only Experiment One	13

Appendix H: (Continued)

Experiment Two: Full Document

	Prep Time			
Group	Hours	Mins	Secs	Score
Full Document	0	48	28	11
Full Document	0	53	5	13
Full Document	0	42	32	14
Full Document	0	18	27	9
Full Document	0	35	38	10
Full Document	0	53	36	10
Full Document	0	35	55	12
Full Document	0	34	29	7
Full Document	0	56	35	10
Full Document	1	0	2	9
Full Document	0	34	18	14
Full Document	0	47	11	7
Full Document	0	27	18	9
Full Document	0	28	3	11
Full Document	0	46	38	17
Full Document	0	16	53	11
Full Document	0	47	27	11
Full Document	0	35	33	13
Full Document	0	38	15	11
Full Document	0	22	57	12

Experiment Two: Context/Keyword Summary

	Prep Time			
Group	Hours	Mins	Secs	Score
Context/Keyword Summary Experiment Two	0	23	32	13
Context/Keyword Summary Experiment Two	0	12	31	16
Context/Keyword Summary Experiment Two	0	32	4	8
Context/Keyword Summary Experiment Two	0	15	56	17
Context/Keyword Summary Experiment Two	0	7	57	11
Context/Keyword Summary Experiment Two	0	42	16	11
Context/Keyword Summary Experiment Two	0	22	50	14
Context/Keyword Summary Experiment Two	0	38	48	10
Context/Keyword Summary Experiment Two	0	28	18	9
Context/Keyword Summary Experiment Two	0	11	25	11
Context/Keyword Summary Experiment Two	0	46	39	14
Context/Keyword Summary Experiment Two	0	22	40	13
Context/Keyword Summary Experiment Two	0	25	28	12
Context/Keyword Summary Experiment Two	0	17	52	10
Context/Keyword Summary Experiment Two	0	42	4	13
Context/Keyword Summary Experiment Two	1	6	43	16
Context/Keyword Summary Experiment Two	0	19	12	13
Context/Keyword Summary Experiment Two	0	37	18	11
Context/Keyword Summary Experiment Two	0	52	4	12
Context/Keyword Summary Experiment Two	0	30	7	14

Appendix H: (Continued)

Experiment Two: Keyword Summary

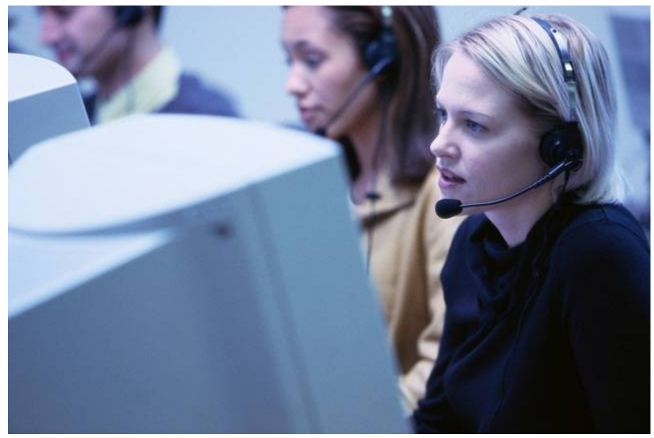
	Prep Time			
Group	Hours	Mins	Secs	Score
Keyword Summary Experiment Two	0	7	11	5
Keyword Summary Experiment Two	0	18	41	9
Keyword Summary Experiment Two	0	37	9	9
Keyword Summary Experiment Two	0	34	24	14
Keyword Summary Experiment Two	0	21	21	10
Keyword Summary Experiment Two	0	25	50	13
Keyword Summary Experiment Two	0	21	45	11
Keyword Summary Experiment Two	0	17	41	11
Keyword Summary Experiment Two	0	59	42	10
Keyword Summary Experiment Two	0	39	54	10
Keyword Summary Experiment Two	0	20	13	10
Keyword Summary Experiment Two	0	17	4	8
Keyword Summary Experiment Two	0	32	43	12
Keyword Summary Experiment Two	0	42	19	13
Keyword Summary Experiment Two	0	38	17	12
Keyword Summary Experiment Two	0	33	40	11
Keyword Summary Experiment Two	0	22	35	10
Keyword Summary Experiment Two	0	20	0	11
Keyword Summary Experiment Two	0	32	15	13
Keyword Summary Experiment Two	0	18	48	12

Appendix I: Document Screenshots

Screen shot of Context/Keyword Homepage for Experiment One:

COP2551 Intro to OOP

Extra Credit



Instructions:

Thank you for participating in this scientific study. Please read the following document using HighBrow. If you choose to highlight a keyword or key phrase, you **must** also highlight the surrounding text as context for the keyword or key phrase. You may take as much time to read the article as you wish, and you may reread it as often as you wish. Please note that HighBrow has the capability to produce a context summary to aid in your recall of important data, directions for its use and all of the other features of HighBrow may be found in the tutorial at the bottom of the page.

There will be a short test on your understanding of the article. The test questions will not involve specific statistics, but will involve general understanding. As an example, a question may ask "Which group of people suffers greater workplace stress while using computers?", **not** "What is the percentage of percentage of Inuit people experiencing workplace stress while using computers? "

Please take the following tutorial to familiarize yourself with HighBrow's capabilities.

Tutorial on HighBrow 2.0

Reading for extra credit:

- THE EFFECTS OF COMPUTERS ON WORKPLACE STRESS, JOB SECURITY AND WORK INTEREST IN CANADA

Visitor 

.

Appendix I: (Continued)

Screenshot of Keyword Homepage for Experiment One:

COP2551 Intro to OOP

Extra Credit

Instructions:

Thank you for participating in this scientific study. Please read the following document using HighBrow. You may take as much time to read the article as you wish, and you may reread it as often as you wish. Please note that HighBrow has the capability to produce a keyword summary to aid in your recall of important data, directions for its use and all of the other features of HighBrow may be found in the tutorial at the bottom of the page.

There will be a short test on your understanding of the article. The test questions will not involve specific statistics, but will involve general understanding. As an example, a question may ask "Which group of people suffers greater workplace stress while using computers?", **not** "What is the percentage of the Inuit people experiencing workplace stress while using computers? "

Please take the following tutorial to familiarize yourself with HighBrow's capabilities.

Tutorial on HighBrow 2.0

Reading for extra credit:

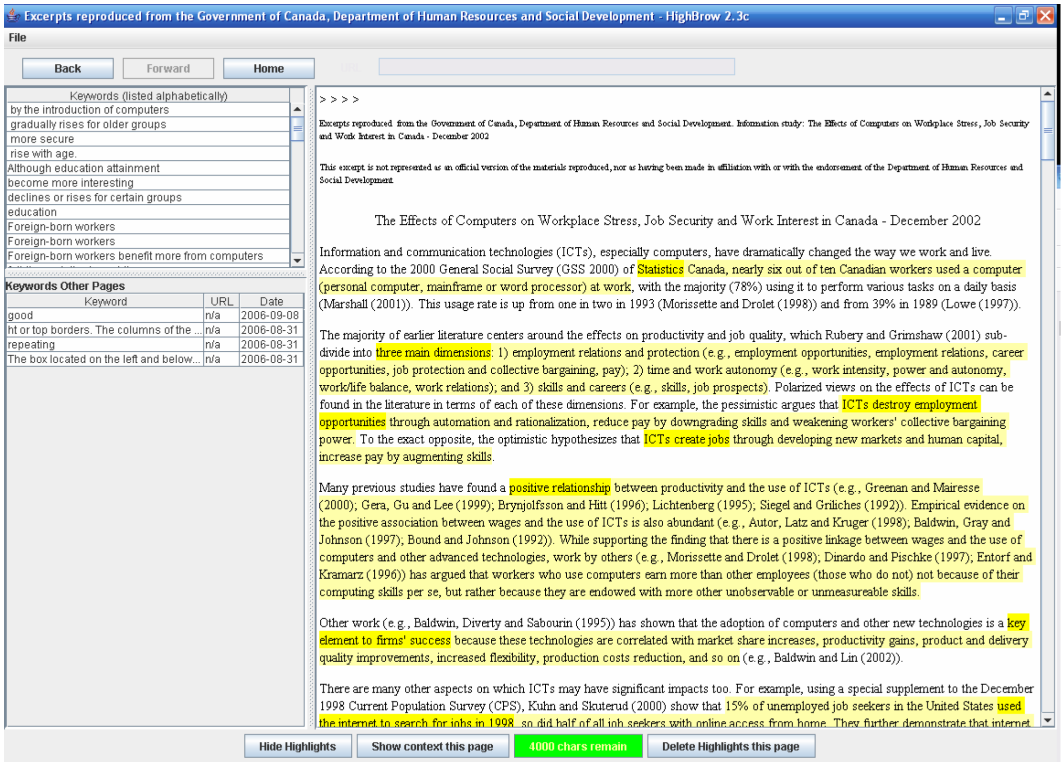
- THE EFFECTS OF COMPUTERS ON WORKPLACE STRESS, JOB SECURITY AND WORK INTEREST IN CANADA

Visitor 

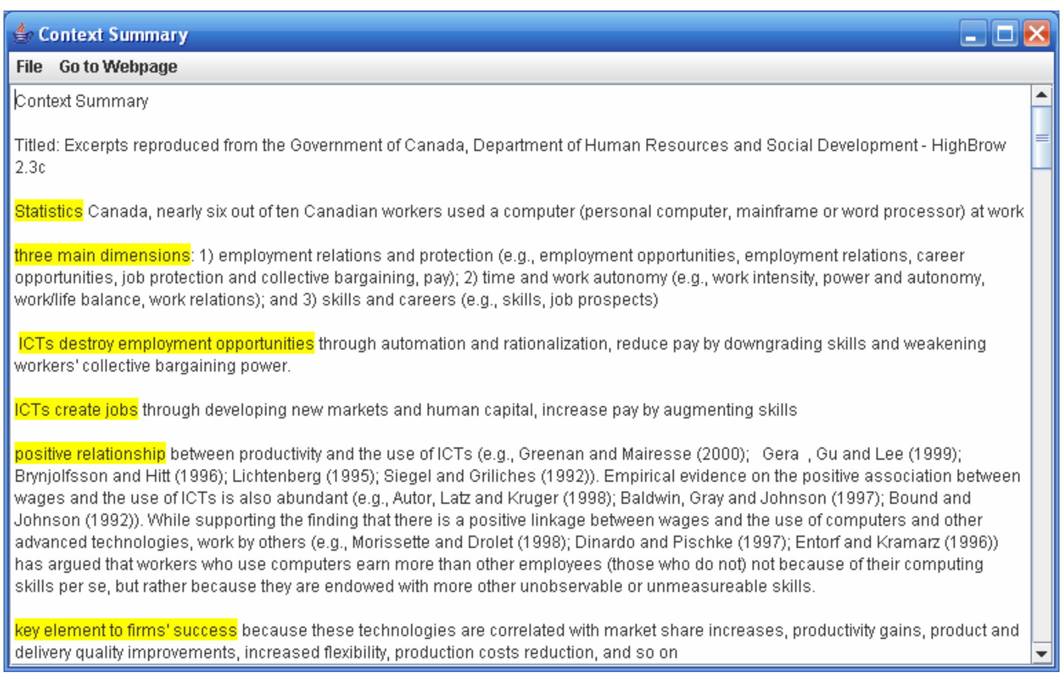
.

Appendix I: (Continued)

Screenshot of MidBrow Interface:

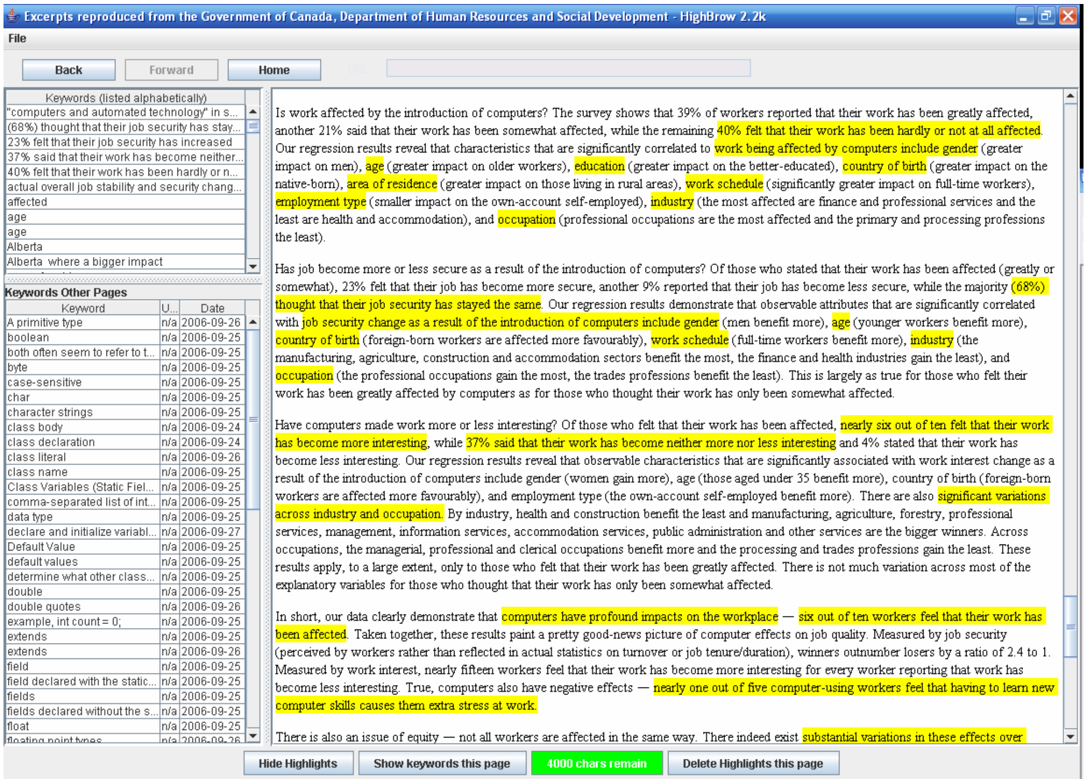


Screen shot of MidBrow Context Summary:

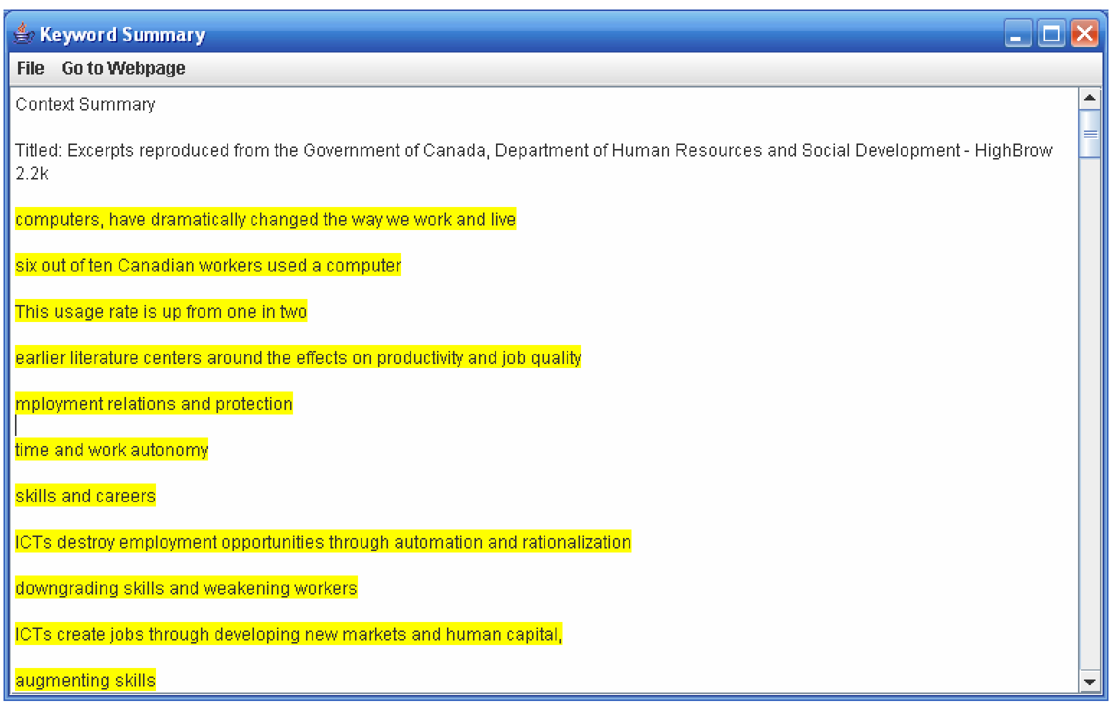


Appendix I: (Continued)

Screenshot of LowBrow Interface:



Screenshot of Keyword summary:



Appendix J: Usability Survey

HighBrow Usability Survey

The feedback you provide about your experience with HighBrow can help us to determine how useful HigBrow was for you. Please take a moment to complete this short survey. For each question please select the most appropriate response. You must respond to every question except the last one, which is optional.

HighBrow userid: ______ (REQUIRED: used for confirmation and credit only, reponses are confidential)

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	n/a
1. HighBrow was easy to install.	☐	☐	☐	☐	☐	☐
2. Once installed, HighBrow was easy to use.	☐	☐	☐	☐	☐	☐
3. Loading web pages was fast.	☐	☐	☐	☐	☐	☐
4. Once a page was loaded, highlighting was fast.	☐	☐	☐	☐	☐	☐
5. It was easy to learn how to use HighBrow.	☐	☐	☐	☐	☐	☐
6. Context Highlighting was beneficial to me.	☐	☐	☐	☐	☐	☐
7. If the HighBrow features were included in my existing browser, I would use them.	☐	☐	☐	☐	☐	☐
8. I liked the layout of the components (indices, navigation, document window, highlighting actions etc.) in HighBrow.	☐	☐	☐	☐	☐	☐
9. My overall experience with HighBrow was positive.	☐	☐	☐	☐	☐	☐

---

10. Capabilities of HighBrow that I liked, disliked, or was neutral or had no opinion.

Ability to:	Like	Dislike	Neutral no opinion
a. Highlight.	☐	☐	☐
b. Modify existing highlights.	☐	☐	☐
c. Delete highlights.	☐	☐	☐
d. Hide/show highlights.	☐	☐	☐
e. Locate keywords for the current page.	☐	☐	☐
f. Locate keywords for other pages.	☐	☐	☐
g. View summary of highlights/context.	☐	☐	☐
h. Print summary.	☐	☐	☐
i. Copy summary to a file.	☐	☐	☐
j. Print original document with highlights.	☐	☐	☐
k. Delete all highlighting on a document.	☐	☐	☐

---

11. (Optional) What, if any, enhancements would you like to see incorporated in future versions of HighBrow?

Submit Survey

Thank you for your response

Appendix K: Institutional Review Board Approval

UNIVERSITY OF
SOUTH FLORIDA

April 24, 2006

Ronald J. Zucker
3237 Hidden Lake Dr. West
Jacksonville, FL 32216

RE: **Requested Revisions** for Application for Initial Review
IRB#: 104622
Title: *HighBrow: A Context Enabled Highlighting Browser*

Dear Mr. Zucker:

On April 21, 2006, the Institutional Review Board (IRB) reviewed your Application for Initial Review. Approval was withheld pending the submission of the following information, which must be submitted no later than two weeks from the date of this letter:

1. Please clarify your affiliation with USF- College of Education or in the College of Engineering or both? Since, the person that conducted the Scientific Review for your study is from the College of Education, but the person that signed off on your study as Departmental Chairperson is from the College of Engineering.
2. There is an inconsistent statement regarding incentives (possible) to participants on page 7 (item #30), page 10 (item 42) and page 12 (item 52). Therefore, please correct and keep the listing of the incentive to be paid to subjects consistent throughout each page of your application. Please provide a copy of your revised application, with all changes highlighted.
3. Please provide a complete copy of your dissertation proposal.
4. Please consider (but note this is not a requirement since your study meets exemption criteria) adding the following USF IRB contact statement to your consent form since USF IRB is also reviewing your study: "If you have any questions about your rights as a person taking part in a research study, you may contact the Division of Research Integrity and Compliance of the University of South Florida at (813) 974-5638."

Please be advised that all of the above information/clarifications must be addressed before final IRB approval can be determined. No research activity including advertising or recruitment can begin until final IRB approval is obtained. If no response is received within 90 days of receipt of this letter, this study will be administratively closed. Should that occur and you wish to pursue this project, you will need to submit a new, revised application for IRB review and approval.

DIVISION OF RESEARCH COMPLIANCE
University of South Florida • 12901 Bruce B. Downs Blvd., MDC035 • Tampa, FL 33612-4799
(813) 974-5638 • FAX (813) 974-5618

Appendix K: (Continued)

UNIVERSITY OF NORTH FLORIDA

ACADEMIC AFFAIRS
4567 St. Johns Bluff Road, South
Jacksonville, Florida 32224-2665
(904) 620-2455 FAX (904) 620-2457

Division of Sponsored Research and Training

MEMORANDUM

TO: Ronald J. Zucker,
CCEC/CIS

FROM: Dr. Kathaleen Bloom, Chair,
UNF Institutional Review Board

DATE: April 6, 2006

RE: Review by the Institutional Review Board IRB #06-047:
"HighBrow: A context enabled highlighting browser"

This is to advise you that your project, "HighBrow: A context enabled highlighting browser," has been reviewed on behalf of the Institutional Review Board and has been declared exempt from further IRB review.

This approval applies to your project in the form and content as submitted to the IRB for review. Any variations or modifications to the approved protocol and/or informed consent forms as they relate to dealing with human subjects must be cleared with the IRB prior to implementing such changes.

Should you have any questions regarding your project or any other IRB issues, please contact Nicole Sayers, Coordinator of Research Compliance, at 620-2498.

nms

c: Dr. Judith Solano, Chair

UNIVERSITY OF
SOUTH FLORIDA

EXEMPTION CERTIFICATION

MEMO: Ronald J. Zucker
3237 Hidden Lake Dr. West
Jacksonville, FL 32216

FROM: Institutional Review Board PGS/amr

SUBJECT: Exemption Certification for Protocol No. 104622

DATE: August 24, 2006

On May 3, 2006 it was determined that your project entitled, **"HighBrow: A Context Enabled Highlighting Browser"** met federal criteria which exempts it from the regulations specified in the Common Rule.

On August 23, 2006, you requested the following changes:

1. Change in the study population:
a. Population for the first study was to come from a course other than the PI's to preserve anonymity. Proposed change is to use students in the PI's own courses COP2551 and COP2551 (2 sections) consisting of a possible 60 students. b. Students in the first study did not follow participation requirements. Students in the context highlighting group were required to highlight context, if they chose to highlight keywords. Many students did not highlight the context, putting them in the control group, not the context group. A general e-mail was sent out to these students informing them of the change, but these students did not highlight the context properly. In addition, approximately 30 students are needed but only 23 participated (out of 23 who remained in the course). A second class had 24 students participate out of 30+ who showed an interest.
c. Extra credit will be given (not to exceed 10 points-one letter grade) to students participating in the experiment. The points will be normalized between the control (keyboard only) and the context (keyboard and context) groups to insure fairness. Students are not required to highlight (in the context group) but are required to highlight context if they choose to highlight. Students who do not wish to participate in the study will be allowed to receive the equivalent extra credit points by writing a paper on the topics covered in the study.

OFFICE OF RESEARCH • DIVISION OF RESEARCH INTEGRITY & COMPLIANCE
INSTITUTIONAL REVIEW BOARDS, FWA No. 00001669
University of South Florida • 12901 Bruce B. Downs Blvd., MDC035 • Tampa, FL 33612-4799
(813) 974-5638 • FAX (813) 974-5618

3. Revisions to the Informed Consent Form to reflect the above referenced changes as well as include USF IRB required statements.
4. Change in the identification of subjects or recruitment procedures: Students will be identified via a userid and will be recruited from classes that the PI is
Based on the above referenced changes (participants no longer remaining anonymous), the status of this study has changed from Exempt Category #2 to Expedited Category #7 with an Informed Consent form. The PI was notified of the status change; and, the approval period start and ending dates were reflected in the letter issued to the PI. The approval period dates for this study are now from 8/23/06 to 8/22/07.

These changes have been noted in the file and do not impact the eligibility for exemption. The study continues to have Exempt Certification. Please remember that any grants connected to this project must be submitted to the Institutional Review Board for review.

Because the study has been certified as exempt, you will not be required to complete continuation or final review reports. However, it is your responsibility to notify the IRB prior to making any changes to the study. Please note that changes made to an exempt protocol may disqualify it from exempt status and may require an expedited or full review.

If you have any questions, please contact the Division of Research Compliance "IRB Administrative Offices" at (813) 974-9343.

pc: Dr. Dewey J. Rundus

UNIVERSITY of NORTH FLORIDA
Division of Sponsored Research and Training
4567 St. Johns Bluff Road South
Jacksonville, FL 32224-2665
904-620-2455 FAX 904-620-2457
Equal Opportunity/Equal Access/Affirmative Action Employer

MEMORANDUM

DATE: August 28, 2006

TO: Ronald J. Zucker,
Computing, Engineering and Construction

FROM: Dr. Kathaleen Bloom, Chair,
UNF Institutional Review Board

RE: Review by the UNF Institutional Review Board IRB#06-047:
"HighBrow: A context enabled highlighting browser"
Amendment Request

This is to advise you that your request to amend your previously-approved protocol, "HighBrow: A context enabled highlighting browser," has been reviewed by the UNF Institutional Review Board and has been approved to include the following:

- Recruitment of subjects from *COP2551 Intro to OOP* and *COP3531 File Structures*
- Revised consent form to include University of South Florida's contact information

This approval applies to your project in the form and content as submitted to the IRB for review. Any variations or modifications to the approved protocol and/or informed consent forms as they relate to dealing with human subjects must be cleared with the IRB prior to implementing such changes. Any unanticipated problems involving risk and any occurrence of serious harm to subjects and others shall be reported promptly to the IRB.

IRB approval is valid for one year. If your project continues for more than one year, you are required to provide an annual status report to the UNF IRB.

Should you have any questions regarding your project or any other IRB issues, please contact Nicole Sayers, Coordinator of Research Compliance, at 620-2498.

Thank you.

c: Dr. Judith Solano, CCEC Chair

About the Author

Ronald Zucker is a Ph.D. candidate in Computer Science and Engineering at the University of South Florida. He has worked in the computer field for twenty years prior to his entry into teaching. Mr. Zucker has taught courses at all levels in the Computer and Information Science curriculum at Troy State University in Montgomery, University of West Florida, Auckland University of Technology, the University of North Florida, and the University of South Florida. Research interests include Human Computer Interaction, Database, and Object Oriented Programming.