THE EFFECT OF PLAYING ONLINE GAMES ON PRODUCTIVITY LEVELS

An Interactive Qualifying Project Report

Submitted to the Faculty

Of the

WORCESTER POLYTECHNIC INSTITUTE

In partial fulfillment of the requirements for the

Degree of Bachelor of Science

by

Michael Anastasia

AND

Jeremiah Chaplin

Date: May 5, 2005

Professor Mark Claypool, Advisor

Abstract

Playing computer games at work is often discouraged, even though there is evidence that work breaks increases productivity and that computer games can be beneficial. We investigated the effects of computer games, particularly short Internet-based games, on productivity, especially when games are played during short breaks. We conducted two studies simulating work in an office environment, and examined the impact of computer game breaks. Although our investigations proved inconclusive, our methodology can be used for future investigations into this subject.

Chapter 1: Introduction

1.1 - Preface

Computer games have existed as a major recreational activity since the introduction of Pong in the 1970's. Since the introduction of the personal computer the availability of electronic entertainment has increased drastically, becoming available in locations where it was previously unaccessible. The sheer number of personal computers in the workplace and presence of games on the World Wide Web have made computer games easily accessible to employee's in the workplace. Like most entertainment, the use of computer games in the workplace is frowned upon by many employers[1], although there is mounting evidence that games can help stimulate thinking and foster social relationships[2]. However, the effects of computer games on productivity have not been as thoroughly explored. This IQP investigated the effects of skill-based computer games, particularly casual online games, on productivity, especially when the games are played during a short work breaks. With computer games on the rise as a recreational activity,

casual games can provide relaxation while keeping a person's mind engaged and focused. This could be especially critical in fields such as customer service or technical support where workers must be ready to assist callers even while waiting long periods for someone to call. This is also important for people working jobs that involve long tedious task such as data entry or secretarial work.

Maximizing workplace productivity is critical to modern businesses. In today's competitive economy, it is understandable that employers want to keep their employees working efficiently. Learning how to best configure and adjust the work environment is critical. The total work environment incorporates many diverse factors that range from the ergonomics of an employee's work space to the length of the lunch break. One factor that has become increasingly relevant is monitoring and controlling Internet access from an employee's computers.

Internet access is critical to many jobs today, but it also presents potential risk to a company. Employees can waste a large amount of time on the Internet if they are not responsible. Each company must decide for itself how closely to regulate access to the Internet. Some companies choose to allow unfettered access; some choose to deny access to certain Web sites. Most often, Web sites with offensive content are filtered out, but other kinds of Web sites are filtered out by some companies. Fearing that their workers will become less productive, some companies have restricted access to web sites that host computer games[3].

Video game players and video game makers do not want to be cut off from one another based on bad reputation alone[4]. While there have been many publicized negative studies regarding computer games, research increasingly shows that video game play

actually fosters cognitive health and improves problem solving skills[5]. While it is not acceptable to play computer games when one is supposed to be doing productive work, it may be acceptable for an employee to be allowed to play games during breaks from work. Game companies have a strong incentives to demonstrate tangible benefits, or at least lack of harm to productivity with scientific data. The goal of this is to study what, if any, impact video game play during breaks have on worker productivity.

1.2 - Productivity

Many of the problems associated with studying productivity stem from the difficulty of defining a quantitative productivity measures. At a very simple level one can just measure the number of tasks a person completes in a period of time and call that the rate of productivity[6]. Unfortunately this does not take into account many other factors that may matter on a management level, such as the number of errors made by workers during the task, the quality of the finished product, or the physical and mental health of their workers. All these factors are important because faulty products, worker exhaustion, and other factors can lead to problems and even extend the amount of total work that must be done due to recall, repair, and rework.

The body of written work on productivity is fairly extensive, ranging from interest articles in business magazines to carefully planned and executed scientific studies. The choice was made to focus on those studies that dealt both with rest breaks and with visual display terminal (VDT) use, primarily computers. Much of the research in this area is concerned with the ergonomics of VDT use, examining musculoskeletal discomfort and stress levels[7]. These same studies also tended to have a fair amount of data on the rate and accuracy of the work their subjects were doing.

The trend that was most prevalent in the studies was that more frequent breaks resulted in higher performance ratings both in accuracy and total production when the total amount of time working and resting was held constant[8]. This implies if two people both work for an hour the person who takes short breaks every 15 minutes will be better off than the person that only takes a long break after 30. This result implies a trend of degradation of performance over time, when a worker starts fresh they have a certain amount of time in which they can maintain a steady work rate before their ability to complete a task begins to decline. After a break worker productivity rate returns to its height and then continues to degrade again as they start work. There is most likely some amount of time that is spent getting back into the work, but as long as the breaks are not clustered too close together, benefits to overall productivity can be observed.

1.3 - Expectations

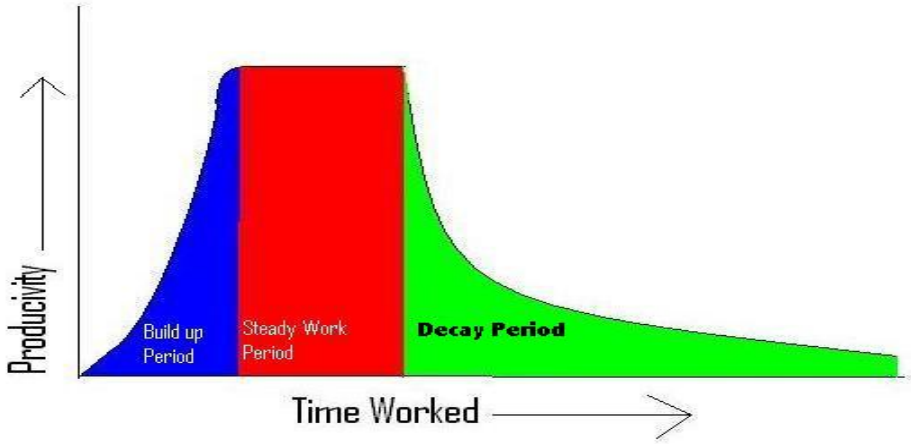


Figure 1.1: Hypothesized Work Cycle, One Iteration

Figure 1.1 displays approximate behavior of a single person's office work productivity over time. First the worker sits down, gets organized, and starts to address

the task at hand. This first stage is called the *build up period*, where productivity slowly rises until it hits some high point. At this point the second stage, or *steady work period*, begins. During this period the worker produces material at a fairly steady rate. After the *steady work period* the worker enters the *decay period*. In the *decay period* the worker's productivity drops off as they become bored, restless, or otherwise tired from their task. The total amount of work they have done during this period can be measured as the area under the curve. The work/rest cycle with a rest break is pictured in figure 1.2. When the worker takes a rest break and then the build up period starts again. While productivity might not rise to the same level it did before the worker took the break, productivity will be higher as they enter the steady work period then it was when the worker left the decay period.

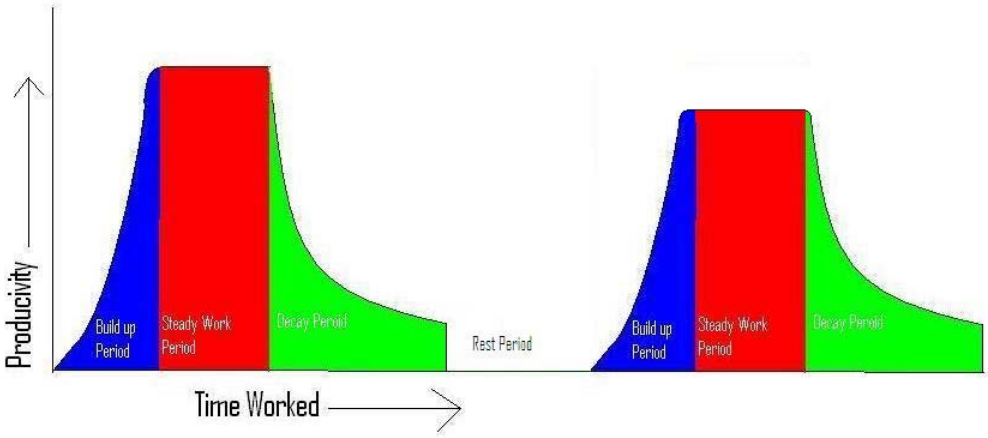


Figure 1.2: Hypothesized Work Cycle, Multiple Iterations with Rest Breaks

1.4 – Methodology

The purpose of this study was to determine whether allowing office workers to play computer games on their work breaks affected productivity after the break.

The study itself was broken into two parts, a preliminary study done to asses the choice of task to see if it produced the theoretical curve of decaying productivity over time and a second user study to determine the effect of different break types. The first study consisted simply of subjects performing a basic task with measurable levels of productivity taken at regular intervals.

The second study measured productivity over time when for one of three different break types; no break at all, a period of doing nothing to represent a basic work break, and a break where the user played a casual computer game before returning to work. The test subjects would perform some measurable task that resembled office work for a given period of time and then would be given a break of some form. They would then return to the task and the change in productivity would be observed.

While the overall methodology of the study appears sound, detailed analysis of the results proved inconclusive. While the preliminary study correctly identified a task that could be used, the nature of the second study meant that incomplete results dominated the data set. Many subjects quit the study process before the break period resulting in a skewed data set. If the same study was performed but the subjects were prevented from producing incomplete results a more definitive answer could be achieved.

Chapter 2: Related Work

There is relatively little research on whether or not gaming breaks affect productivity. The two related areas of research are studies on the topic of work breaks and studies of productivity in general. Studies on productivity provided ideas as to how to design our own study, while the studies on work breaks gave us an idea of how merely not working for a segment would affect our study. This would mean that we could have a clear idea of the effect of the games themselves rather than merely the break.

2.1 - Rest Breaks and Measures of Productivity

The first research area, studies that focused on measuring productivity, gave a provided a background on how one goes about measuring the concept. In *Developer Producitivity Study* the Branham Group studied two different development suites and which one aided productivity more. The study used a wide variety of test for productivity including code size, time, and whether or not it might the requirements set for the specific test. These are solid metrics for measuring how productive an item of interest is. Time and the products ability to meet a given requirement were adapted as metrics to be used in this study.

The second area of research involved the effect of work breaks on productivity. The research indicated that an appropriate work/rest schedule, defined as a period of work and then a period of rest, improves productivity. When workers were asked to engage in computer related tasks for an hour with 10 minutes of break time, the correlation was that the more frequent the break the higher productivity rate. Also of note is the fact that a cognitive task (basic arithmetic) caused discomfort and performance

loss at a much higher rate than a data entry task[9]. This study provides not only a comparison for the results of this study but also is an example of how a study can be constructed to examine how work breaks affect productivity. Other performance factors such as employee burnout and accident rates are also affected by the work/rest schedule that individuals are working under. As workers were asked to work longer shifts the chances of a serious accident occurring or workers reporting excessive fatigue increased[10,11] . This corraborates the idea that work breaks improve producivity since accidents and worker burnount will both detract from the quality of the finished product and potentially hamper the products ability to meet requirements.

In *Compensatory rest breaks for VDT operators* it was found that how rest break is scheduled; whether worker are allowed to select when to rest or if they forced to rest based on a schedule does not noticeably affect performance. The study examined error, correction, and completion rate for a group of 38 users and only found significant variation in how much lower back discomfort experienced by the users, the actual productivity measures were similar between the two groups[12]. This means that the use of scheduled work breaks as opposed to user selected work breaks does not need to be considered as a nuisance factor. The study was designed with this fact in mind and used a scheduled work break for technical simplicity.

2.3 – Effects of Computer Games

Research involving games and their affect on growth and development has been around for some time and there exists evidence that computer games enhance reflexes and eye-hand coordination[13]. The body of work for the effect of computer games is rather small. However, computer games consist of a powerful immersive component,

often considered as the reason that computer games can incite violent behavior in children. However, this same immersion component involves and engages the player's mind, resulting in high level of concentration and focus[14]. If this component is helpful for learning and students, perhaps it will also help maintain concentration and focus during rest breaks such that transitioning back to a work schedule is not quite so difficult. This immersion can cause problems as workers heavily engrossed in a game might not easily return to their work if the game has captured their attention too deeply. Computer games are enough of a problem that many companies block Web sites or pay for software that will prevent their employee's from accessing game sites. One article estimated the loss caused by workers playing games to go as high as $50 billion[15]. While people should not be playing games when they should be working; games during rest breaks might serve as an effective way to maintain worker focus even while they are not on the job.

Chapter 3: Pilot Study

In order to determine the effect that games had on productivity, we first needed to create a baseline measure of a person's productivity. Since our primary focus was people doing office-type work, sitting at desk and working on a computer, a task intellectually similar to the standard office work but simple enough for anyone to understand would provide the good baseline from which to determine what effect computer games were having on subject productivity. The first priority was to confirm that productivity followed the expected behavior and to test that the task that was given to the subjects in the study was analogous to working at an office job. To this end we chose to conduct a small pilot study that would both confirm the expected behavior and solidify the precise methodology of how to administer a larger study to determine the impact of games on the task.

3.1 Task Selection

The question of what task to use was an important one and proposed ideas for the task ranged from a series of simple math problems to data entry. Trying a series of math problems on a few volunteers revealed that basic arithmetic was too boring to produce the expected behavior; people became bored and withdrew from participation two minutes into the study, far before their performance started to decline in any meaningful way. Also, there could be a large and unpredictable variance in performance between subjects due to the variance in the population regarding their arithmetic competence. The task was switched from arithmetic to transcription; the subject would attempt to copy a document by typing it word for word. Five samples, each of a different style of writing were selected and given to each subject for transcription. This task was much better

suited to our needs since it had a smaller range of variance than the arithmetic and the content of the written material kept the subjects interested and working for longer periods of time. However, five pages turned out to be too much material to cover in a reasonable amount of time. Given that subjects of the study would be uncompensated volunteers, the study was constrained to a maximum duration of one hour. Five pages took longer than one hour on average so the number of pages that each subject would transcribe was dropped to three. The three pages were chosen at random for each user from the initial five.

The transcribed document was merged with the original, and the number of errors made by the subject in transcribing the document was counted. The measure of productivity included both time and accuracy. Accuracy was measured with Microsoft Word's merge and compare function; which highlights any discrepancies, grouping nearby errors together into one error. Capitalization, misspellings, and any other errors on a character by character level were all counted as errors. Word also highlights as errors formatting differences and any missing or added line breaks, but we did not include these errors in the count. Time taken per page was measured by the user marking the time they finished a particular page. Using the time markers (which were accurate to the minute) and the error count, a measure of subject productivity over time could be calculated.

3.2 Methodology

The participants in the pilot study were friends of the authors who were willing to volunteer their time. The instructions given to each subject were as follows:

1 . You will be asked to copy three pages of text

2 . Copy each page as quickly and accurately as possible

3 . After each page, please indicate each of the following on a scale of 1-10

- How interested are you in the content of the article you copied? (One means not at all interested, ten means very interested)
- How bored are you with the task of transcribing documents? (One is not bored at all, ten is very bored)
- How good do you think your transcription is? (One indicates many errors, ten indicates a perfect copy)

4 . Then check in with the study administrator, who will record the time you took on this page and tell you what page to copy next.

Subjects took the study at the computer most convenient to them. In some cases, this was the subject's personal computer; in others it was a laptop provided by the authors. Regardless of where the study was conducted, the on screen environment was the same for each subject. Two instances of Microsoft Word were open, each filling one half of the screen. One half showed the document to be transcribed; the other was for the subject to type the transcription. The subject noted the three qualitative measures at the bottom of each transcribed page. The subject's next page was chosen randomly from the pages not yet typed. After the subject had completed three pages, the number of errors on each page was computed by using Microsoft Word's compare and merge feature to

compare the original document and the subject's transcription.

The five documents selected for this task were

An excerpt from the short story "The lady and the tiger"

An excerpt from a report about the rivalry between the Boston Red Sox and the New York Yankees

An excerpt from The Odyssey

An excerpt from the Wikipedia entry on the British railroad system

An excerpt from a scientific report on the merit of sealing rooms with duct tape to protect from chemical attacks

The full text of each document used is in Appendix A.

3.3 Results

Five subjects transcribed 3 pages of text, each page randomly selected without replacement from the five samples. Due to the nature of the randomization, there are two ways to evaluate the error numbers. Reporting could be grouped together by source document or by the order in which they were seen by a specific user. Since the concern of this study is the effect of breaks on a single person over the course of a prolonged period of work, the data gives analysis is most meaningful when organized in chronological order of transcription, regardless of which documents the subject transcribed.

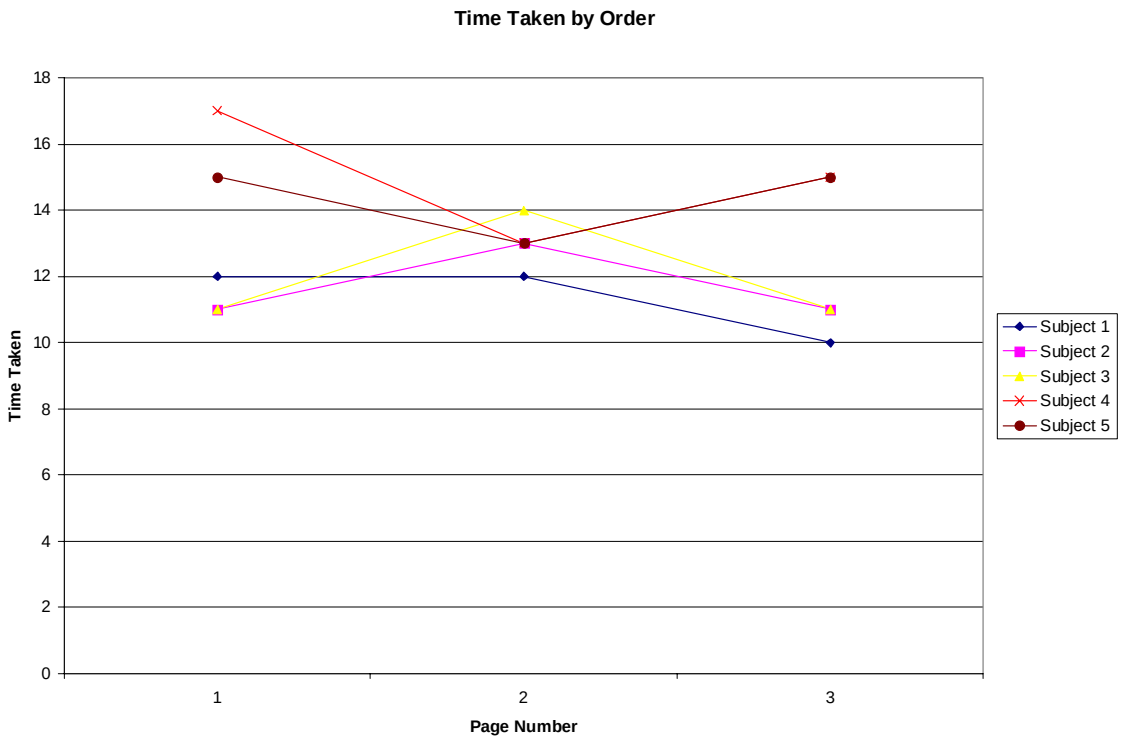


Figure 3.1: Pilot Study Time per Page

Figure 3.1 plots the amount of time in minutes each subject took to complete each page on the y axis against the page number on the x axis. Each data set in the figure corresponds to a unique subject in the pilot study. The time a given user took to complete each page was consistent, although it does show a slight downward trend as page number increases. Productivity is often measured in terms of the number of work units completed divided by the time taken to complete them. Since most users took between 11 and 15 minutes per page, without a noticeable pattern of change over time, the time component in the units/time equation is the same for all subjects and thus can be treated as 1 unit. Since time was fairly uniform across all subjects we can look at the number of errors as an absolute measure of worker performance in this study.

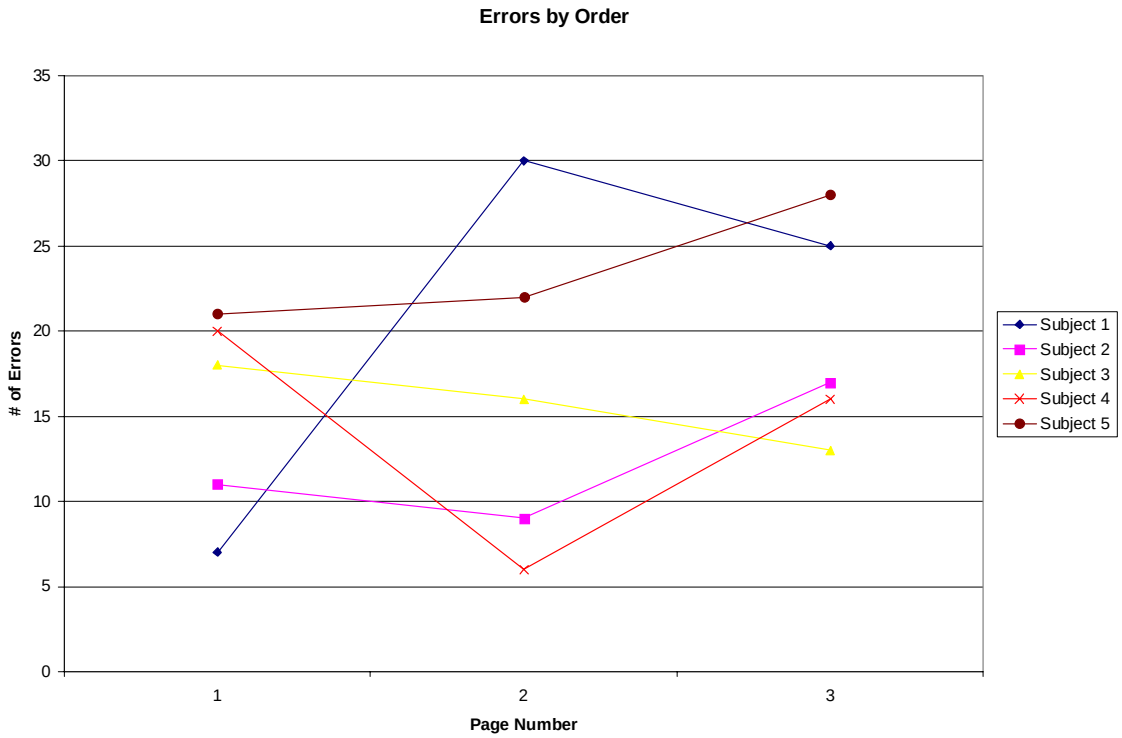


Figure 3.2 Pilot Study Errors per Page

Figure 3.2 plots the number of errors each subject made on the y axis against the page number on the x axis. Each data set in the figure corresponds to a unique subject in the pilot study. Errors themselves remain fairly consistent until the end, when they tend to trend upwards. Interestingly enough, most subjects made fewer errors between the first and second pages, implying that perhaps the amount of time required to get into the highest level of productivity is longer than originally thought. The upward trend follows our hypothesis, especially when one considers that the time spent per page remained roughly constant through the study. The small difference between starting and ending errors is of some concern, but for this most part this graph shows that transcription may be an effective task for inducing increasing errors in a short amount of time.

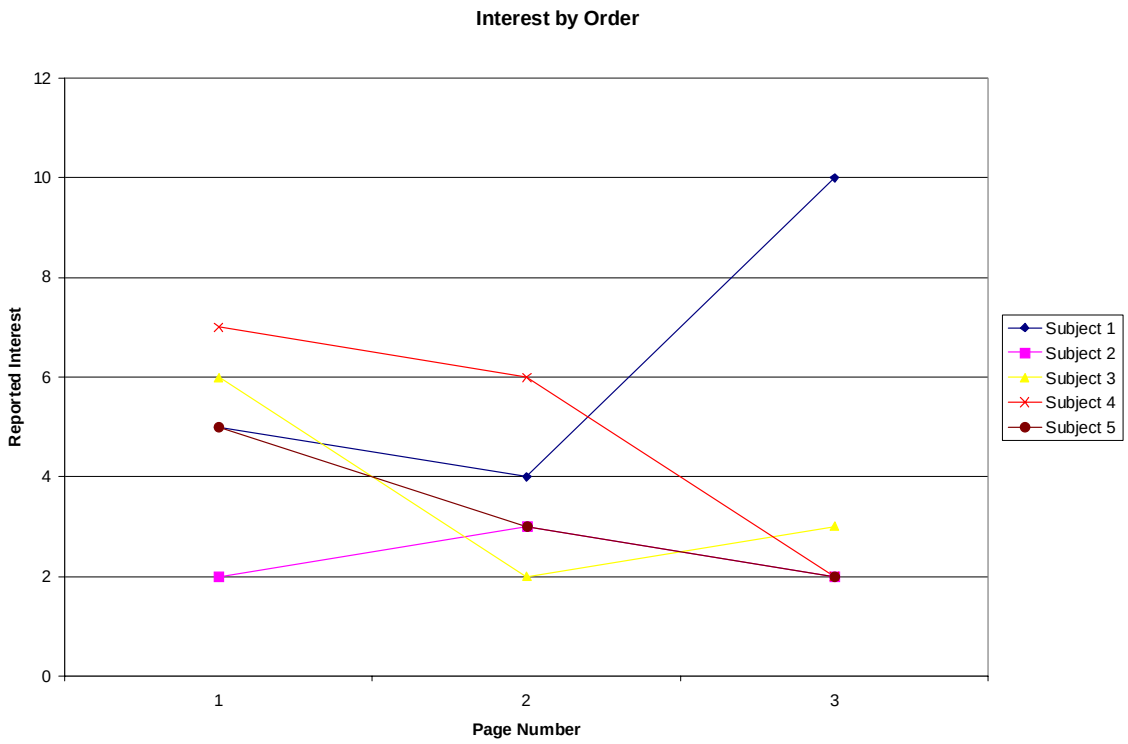


Figure 3.3: Pilot Study Interest per Page

Figure 3.3 plots the interest response from each subject on the y axis against the page number on the x axis. Each data set in the figure corresponds to a unique subject in the pilot study. The question asked was "How interested are you in the content (subject matter) of this page? One represents no interest and ten represents high interest." Interest declines slightly from page one to page two, but between pages two and three the trend becomes inconsistent. It is likely that both of these effects are coincidental because the pages are sorted by order copied, the interest ratings for a certain page number here are related to the same document. Personal taste also impacts these ratings, and interest in different topics varies from person to person.

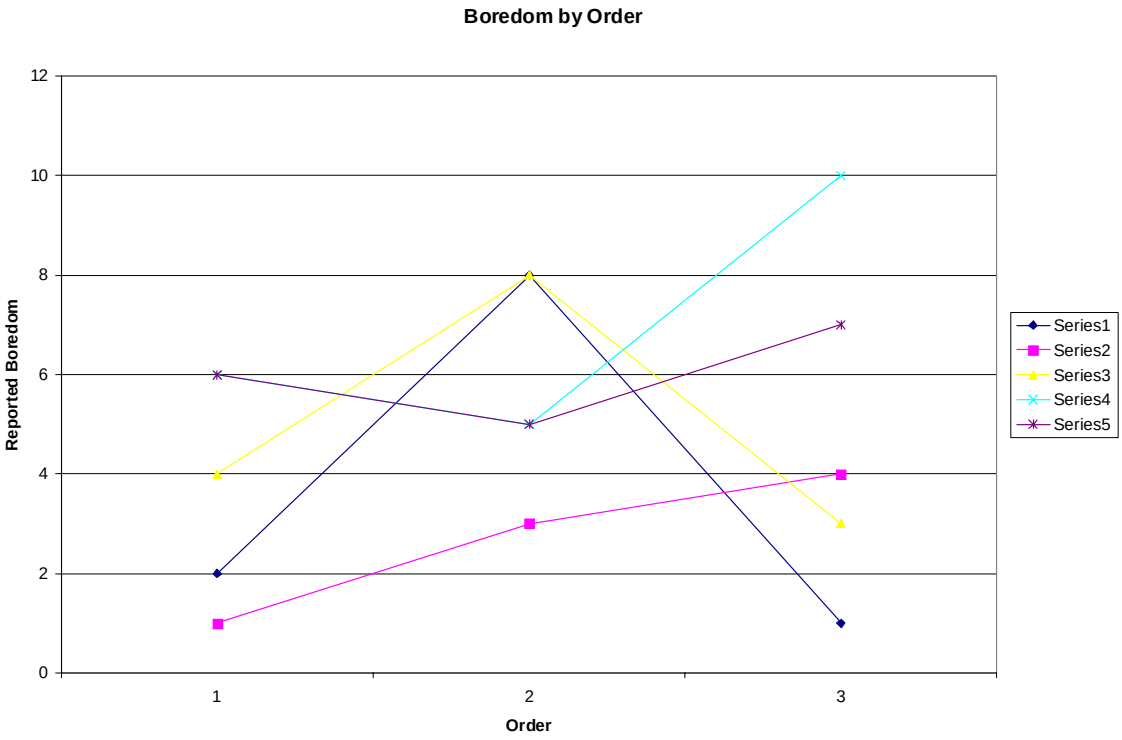


Figure 3.4: Pilot Study Boredom per Page

Figure 3.4 plots the boredom response from each subject on the y axis against the page number on the x axis. Each data set in the figure corresponds to a unique subject in the pilot study. The question asked was "How bored are you with the task itself, where one means not at all bored and ten is very bored?" Reported boredom shows no clear trend across the different subjects; some become more bored over time, some less, and some remain the same. This is unexpected because one would expect a repetitive task to cause to an increase in boredom.

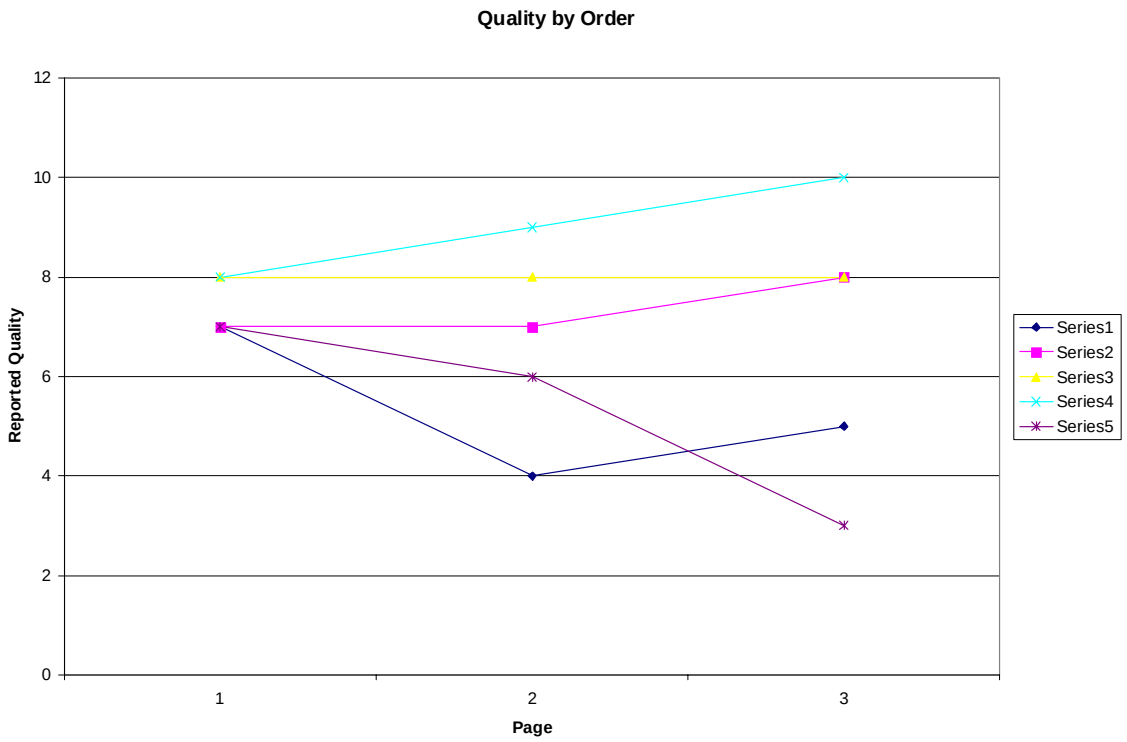


Figure 3.5: Pilot Study Quality per Page

Figure 3.5 plots the quality response from each subject made on the y axis against the page number on the x axis. Each data set in the figure corresponds to a unique subject in the pilot study. The question asked was "How well do you think you copied the document, where a zero is many errors and ten represents a perfect copy?" Self-assessment of quality shows little to no change over time, implying that users think that they are performing at the same level even though performance as measured by error count is actually varying.

3.4 Analysis

All subjects, with the exception of subject 3, showed an increase in the number of errors over time. Some users actually declined in errors on their second page, implying that they were getting better at their task with time. This implies that in the time it took

them to transcribe one page they were in the build up period (see Figure 1.1), which might take more time than was expected. Also of note is subject one, who appears to hit his threshold early then catches a second wind on the third page.

The results of the data analysis were not as strong as hoped; although the subjects did become slightly more productive on page two, and then less productive again on page three, not all followed this trend. An ideal task for this research should reproduce the expected behavior from Figure 1.1 in a time span of roughly thirty minutes to an hour. Examining Figure 3.1, several of the subjects do follow the pattern of improvement, stable performance, and then decay. However, only three of the five subjects showed exactly this trend and the other two fell well outside expectations. Subject 1 seemed to enter the decay period almost immediately, and subject three did not appear to enter the decay period during the time of the pilot study.

Despite the result, the trend does lend some support to our hypothesis that the task of written transcription does in fact show our hypothetical productivity decay. The small number of subjects means that the findings lack any real statistical significance and the presence of outliers could have caused the low adherence to the expected behavior. The initial plan to have all of subjects transcribe all of the pages was quickly discarded when it was realized that five pages of transcription was roughly two hours of work, far too much to ask of people if we wanted a significant number of people to participate. The choice to give every subject a random order of three pages was made consistent analysis much more difficult. The fact that each user had a different set of parameters under which they worked and the noise generated from that threw certainty of the results even farther off, enough that it is interesting that the expected behavior showed up at all.

Despite the inconsistencies created by the randomization of pages, most users showed a productivity increase followed by decay. It appeared that with some changes to the methodology transcription is an acceptable task to use in a complete study of the effects of computer games on productivity levels.

Chapter 4: User Study

We could have next designed and conducted such a large study to determine more conclusively if transcription was the correct task for our study. However, due to the limited time available to us, we chose to proceed with a study of productivity and computer games.

The pilot study was constructed with the question “Is transcription an appropriate task for a study of productivity and computer games?” in mind. Since transcription seemed an appropriate choice for analyzing productivity levels the decision was made to move ahead with the full study. If the for some unforeseen reason the task turned out to be unsuitable, the result would be the same as that of a large version of the pilot study: that is, our approach would have proved incorrect. On the other hand, if the task was appropriate, then at the completion of the full study we would have all the data we need to examine the core area of our investigation of computer games and their impact on productivity.

4.1 Design

A number of elements of the pilot study were kept exactly the same as in the pilot study. We continued using transcription for our task; transcription tests are not difficult to implement, most potential subjects have all the necessary skills, and other research in this field has used it in the past. Data gathered were the same three qualitative data after each page, Boredom, Interest, and Quality, and the same quantitative data, time per page and the direct text input for each page. The instructions given to each subject were more or less unchanged, although they were repeated at the beginning of each page for clarity. We continued to allow subjects to work on the computer of their choosing. This freed

them to participate at their convenience and also removed any alteration of the data derived from subject unfamiliarity with the computer.

The design of this new study encompassed two main thrusts. First, we adjusted and tuned the pilot study based on lessons learned from running it the first time. Our goal was to have better control over the environment and to reduce external sources of variation, while simultaneously adapting the study for a larger sample. Second, we added rest breaks of different kinds to the study so we could measure their impact on productivity.

4.2 Interface Changes

We made a number of changes to the design and execution of the full study based on our experience with the pilot. The first, and largest, change from the pilot study was that the full study would be conducted over the Internet. This way, the interface each subject saw on his screen was exactly the same from one user to the next. By placing the source documents on the screen as pictures rather than text, we thwarted any attempt to 'cheat' using copy and paste. Using Web pages also allowed us to automate data collection. That was critical; gathering the data and computing the results by hand was practical with only five subjects, but with the larger sample, the tabulations would have required more time than was available.

4.3 Document Selection and Preparation

A source of variation identified by the pilot study was the change of source documents from page to page. Since documents differ in style, character count, diction, and other factors that would obscure the results we needed to find some way to

compensate for them. This was accomplished by drawing all the transcriptions from the same source document, assuming that consistent style and diction would be maintained throughout the entire document. We selected a single article as the source for every entry in the study. Each page for transcription was the next section of the article. Even though no two pages of the article were exactly identical, because they came from the same source, they were similar in tone, style, and subject matter. This was important for eliminating any variance due to the nature of the material being copied. If every subject typed the same pages in the same order as each other subject and every page is similar to (i.e. from the same source as) each other page, comparisons across different subjects and across different pages became much more consistent.

We also chose to break the document up into smaller pieces. We divided the three page article into six half-page entries. Where a subject in the first study would type a full printed page, a subject in the second study would type roughly half of a printed page. As with the first study, the subject gave qualitative assessments of interest, boredom, and quality after each of these pages. The result of this change is that the data, both qualitative and quantitative, is recorded at twice the resolution of the first study, or roughly once every five to seven minutes.

We chose a news article about the federal government as our source document. This article was appropriate for a number of reasons. Firstly, its content was likely of some interest to our subjects, and would be familiar to many of them. We were careful to choose an article that did not show a large political bias towards one opinion or another. Another positive feature of a news article was that it had many quotations and proper nouns; we hypothesized that these are harder to transcribe perfectly resulting in more

errors per page to begin with and hopefully a more easily observed decay period. The original text of the article, along with the dividers for each entry, are in Appendix B.

4.4 Productivity Measurement

Rather than manually counting errors using Microsoft Word's compare feature, we automated the error counting for this study. The UNIX diff command compares documents line-by-line. We wanted to compare character-by-character, so our error-counting tool inserted a new line after each character in the source and input documents and then called diff on the two documents. The number of incorrect characters could then be determined. In this method, any mistake on a keystroke counted as one error. Capitalization mistakes, inserted or omitted characters, and inserted or omitted whitespace all counted as errors. In order to learn more about the specific kinds of errors users made, we also ran the input documents through the UNIX spell command to catch dictionary spelling errors. If the patterns of errors were different by this measure than by strict 'incorrect keystrokes', we determine what percentage of errors are spelling mistakes and what percentage are repeated words, misread words, and miscopied punctuation and/or spacing.

4.5 Incentive

In order to encourage subjects to finish the study in its entirety, we decided that some incentive would be needed. Ideally, we would have compensated all our participants, but resources to do so were not available. We offered two twenty-five dollar gift certificates, to be raffled to people who completed the study. We advertised our study by email and printed flyers to the Computer Science department at WPI, to several IMGD classes at WPI, and to the WPI Science Fiction Society. A copy of the flyer is in

Appendix C. These groups were targeted because the authors and the advisor of this project are members of the groups and had easy access to the students.

4.6 Rest Breaks

The last change we made to the study design was to actually incorporate breaks into the study. Our new article was broken into six segments, so we placed the break after the third segment. One third of the participants would get no break and would continue directly to segment four. One third would be instructed to take a five minute break to get a glass of water and relax, and the final third would play a video game for five minutes before continuing. These three groups were designated A, B, and C, respectively, and are also called the no-break, rest-break, and game-break groups. Comparing the performance of subjects in these three groups allows us to draw conclusions about the effects of computer game breaks on productivity.

4.7 Implementation

In order to create a standardized environment that could be accessed from any computer, we implemented the study as a series of Web pages. After exploring our options for Web scripting, we chose to use Perl CGI scripts to control the flow of the study, primarily because it was supported by the WPI Web server. The study was initialized and the subject placed randomly in one of the subject groups by a Perl script we made ourselves. WPI provided webform.pl, a script for gathering Web input and writing it out to text files. All of the transcription and data gathering pages in our study were powered by this script.

For the game, we searched for a well-known game implemented in Java. Many Web games are written in Java and we expected that we could find a simple, well-known

game that could launch from the Web. The first promising game we found was *TubeBlazer*[16], a simple reflex game that launched using Java Web Start. Unfortunately, Java Web Start is not very common, and it is not installed in WPI's computer labs. We decided that we needed to find a different game that would run as a Java applet, which more subjects could actually run without an additional download. In the end, we selected an open source Pac-Man clone to use for our study[17].

The subjects went through the series of Web pages in order, the sequence as follows:

I. Background information and terms and conditions (at the completion of this stage the subject was randomly placed in the no break, rest break, or game break groups)
II. Input demographic information
III. Instructions
IV. First transcription
V. First evaluation
VI. Second transcription
VII. Second evaluation
VIII. Third transcription
IX. Third evaluation
X. Break period. (If the subject was in the no break group they skipped this step and proceeded as normal. If the subject was in the game break group they played the Pac-man game for five minutes, then proceeded. If the subject was in the rest break group they took a rest break for five minutes before proceeding)

XI. Fourth transcription

XII. Fourth evaluation

XIII. Fifth transcription

XIV. Fifth evaluation

XV. Sixth transcription

XVI. Sixth evaluation

Screen shots of the Web interface are in Appendix D.

We made number of considerations in order to compensate for the online, unregulated nature of the study. At the start of the study, subjects read the instructions for the study and indicated their willingness to participate. They were also told that two gift certificate to Best Buy would raffled to people who finished the study. The instructions for each section were repeated at the top of each page, along with a request not to use the browser 'back' button, as this would disrupt our data collection script. To prevent subjects from simply using copy/paste functions to transcribe the text perfectly, we converted the text of the document into an image file for display. Subjects were allowed to quit the study at any time by closing their browser window.

Two aspects of the implementation deviated from the original design of the study. Firstly, we were unable to implement automatic advancement past the break after a certain time, so we instead got the system time from the computer, and printed on the screen an instruction telling the subject at what time to stop resting or playing and to click next. Subjects overwhelmingly followed this instruction correctly. Secondly, the fifth transcription was mistakenly typed in as a second copy of the first page. Some subjects saw this repeat and quit the study, assuming that it would go on forever. Others either

ignored the repetition and continued or did not notice.

During the first day of the study there was a Web error reported by several members of the no break group. The error was corrected quickly, but some subjects in the no break group were unable to complete the study due to the technical error.

Chapter 5: User Study Results and Analysis

Once the study had been conducted and the data imported into a spreadsheet the process of interpreting the data came forward. Initial observation showed that the data itself was somewhat sporadic; there was clear decrease in the number of subjects who started the study as well as some data points that were significantly off from the others. The first issue that was dealt with was determining the cause of some of the more anomalous results and trying to create a picture of what actually happened during the study.

Thirty-eight subjects began the study, ten in Group A (no break), fourteen in Group B (rest break), and fourteen in Group C (video game break). Nineteen finished the study in its entirety, three in Group A, seven in Group B, and nine in Group C. The decline over time is examined closely in Figure 5.1.

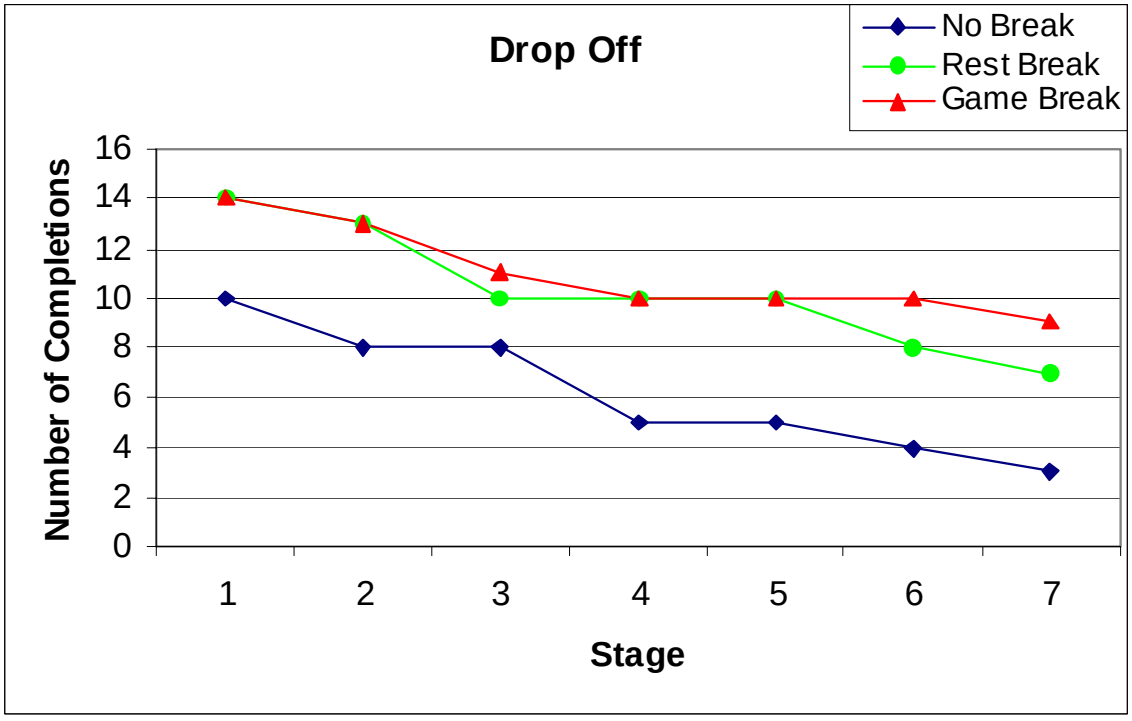


Figure 5.1: User Study Page Completions by Group

Figure 5.1 plots the number of subjects that completed each page on the y axis against the entry number on the x axis. Entry number four corresponds to the break, and entry numbers five, six and seven correspond to pages four, five, and six. Each data set in the figure corresponds to one of the Groups, counting all responses.

A combination of technical glitches and subject's free will resulted in many subjects quitting the study partway through. In addition, because group assignment was random and made no attempt to even out the groups if one was higher than the others, group A (no break) started off smaller than the other two groups.

We believe that people left the study roughly when they would have otherwise entered the decay period of our initial hypothesis. That aside, the dropout rate for the rest break group and the game break stay fairly constant, although the game break is slightly stronger in terms of subject retention, especially at the tail end of the study. Also of note is that in the no break group there was a sharp decline at stage four, which was the break itself. This implies that by the time the break was given many subjects would have dropped and that the presence of a break retarded the rate of subject dropout. The dropout rate supports the hypothesis and also shows that game breaks hold the interest of more subjects over the long term.

5.2 Data culling

Interpreting the data generated by the study proved fairly difficult and a number of techniques were used in order to remove the outliers and remove the noise from the data set. Due to the high number of outliers in the data only the median is examined in detail here. The data was also culled due to some incomplete or erroneous results, subjects who did not reach page 4 or had error rates an order of magnitude above the

median were removed from the data before analysis began. The reason the high errors were culled was because they were fairly evenly distributed among the groups and the cause of these errors was actually subjects who did not complete the study but clicked through all the pages. This meant that they were submitting empty text boxes which registered each character in the passage as a keystroke error. These results were culled out and counted as incompletes. Even this reduced data set had several large outliers; the average of measured values demonstrated no statistical significance. This section only includes the best representation of the data we collected, although more graphs for each technique are included along with brief explanation in Appendix E. In our analysis of any given statistic, we considered both the median value and the mean value from each of groups A, B and C. Although most of the outliers were removed, the data was still spread fairly sporadically and so the median was deemed to more clearly measure productivity.

5.3 Results

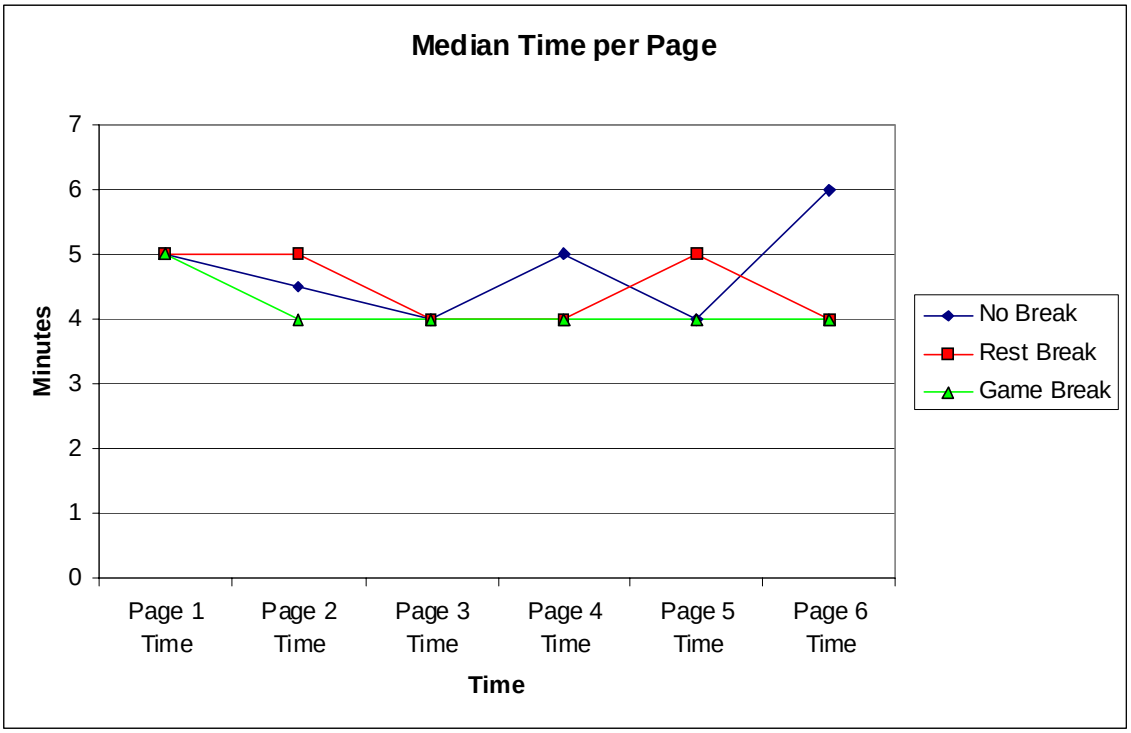


Figure 5.2: Median time per page.

Figure 5.2 plots the median amount of time taken by subjects in each group on the y axis against the entry number on the x axis. Each data set in the figure corresponds to one of the Groups, counting all responses in the reduced data set. An analysis of the time per page shows behavior very similar to what was seen in the pilot study; the time taken on any given page stayed fairly constant over the entire study. Everyone took about 4-5 minutes per page with an outlier of 6 minutes on the last page of the no break subject group. As in the pilot study, since time per entry remains flat over all groups and throughout time, the time component in the errors/time=productivity equation is the same for all subjects and thus time can be treated as 1.

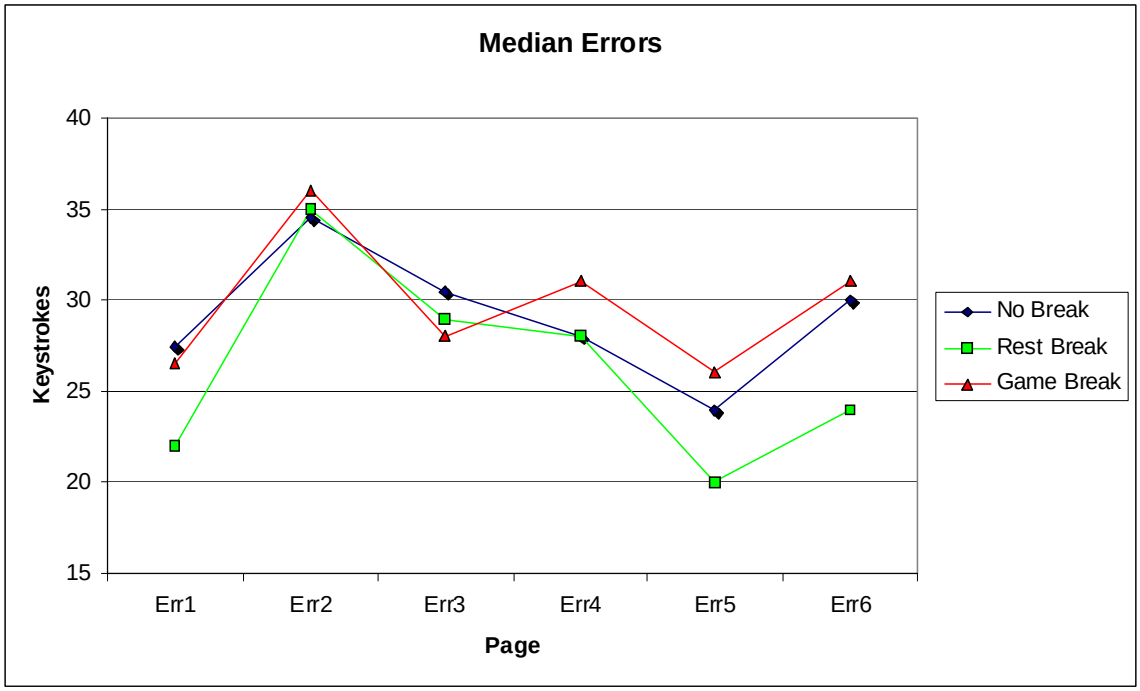


Figure 5.3: Median Errors per Page

Figure 5.3 plots the median number of errors subjects made in each group in each entry on the y axis against the entry number on the x axis. Each data set in the figure corresponds to one of the groups, counting all responses in the reduced data set. This graph represents the median for the number of errors for all subjects who completed at least the fourth transcription. An error was defined as a character which differed from the article; including any difference in white space. A word with multiple misplaced letters or spelling mistakes counted as one error for each misplaced or missing character.

The data varies between pages but not outside of 20-35 errors per page. There is a slight trend upwards between page 1 and 6, but only by a margin of 3-5 errors. The large jump in errors in page 2 is the most obvious feature of this graph, albeit a perplexing one. Page 2 seems rather early for people to already be suffering performance degradation since by that time they would only have been typing for 5-7 minutes on average. The simplest explanation is that for some reason page 2 was more difficult to

transcribe than the other pages, resulting in more errors overall.

The critical junction lies between pages 3 and 4, the pages which straddle the break itself. The difference in errors between the two pages is almost negligible within all the subject groups. The game break actually sees a minor increase in errors after the break was administered while the other two groups see a minor decrease in their errors, including the group which received no break.

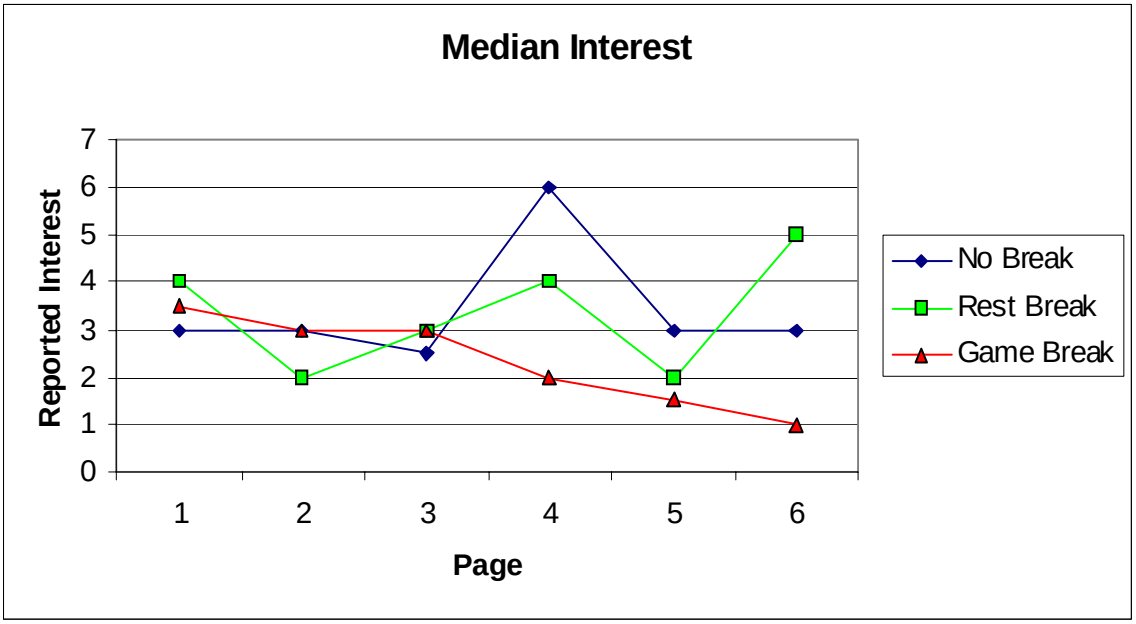


Figure 5.4: User Study Interest per Page

Figure 5.4 plots the median interest response from each subject on the y axis against the page number on the x axis. Each data set in the figure corresponds to one of the Groups, counting all responses in the reduced data set. The question asked was "How interested are you in the content (subject matter) of this page? Zero represents no interest and ten represents high interest." Interest varied quite a bit over the course of the study, and did not show a clear trend over time. It is surprising that ratings of interest varied as much as they did, given that every entry came from the same article.

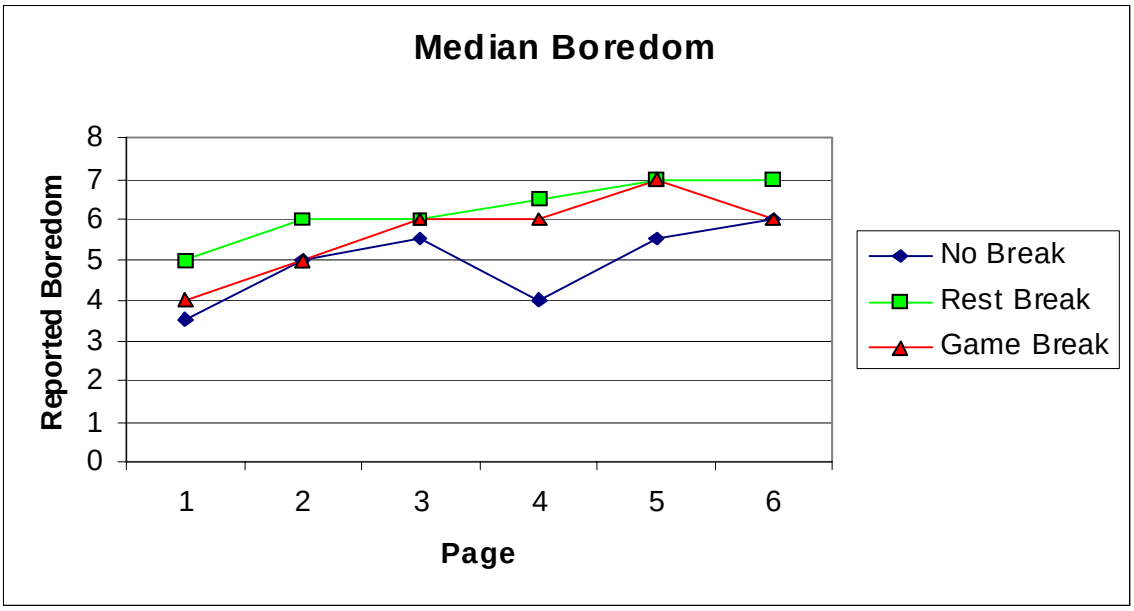


Figure 5.5: Pilot Study Boredom per Page

Figure 5.5 plots the median boredom response from each subject made on the y axis against the page number on the x axis. Each data set in the figure corresponds to a unique subject in the pilot study. The question asked was "How bored are you with the task itself, where zero means not at all bored and ten is very bored?" As expected, boredom showed an upward trend, increasing by 3-4 points on average over the course of the entire study.

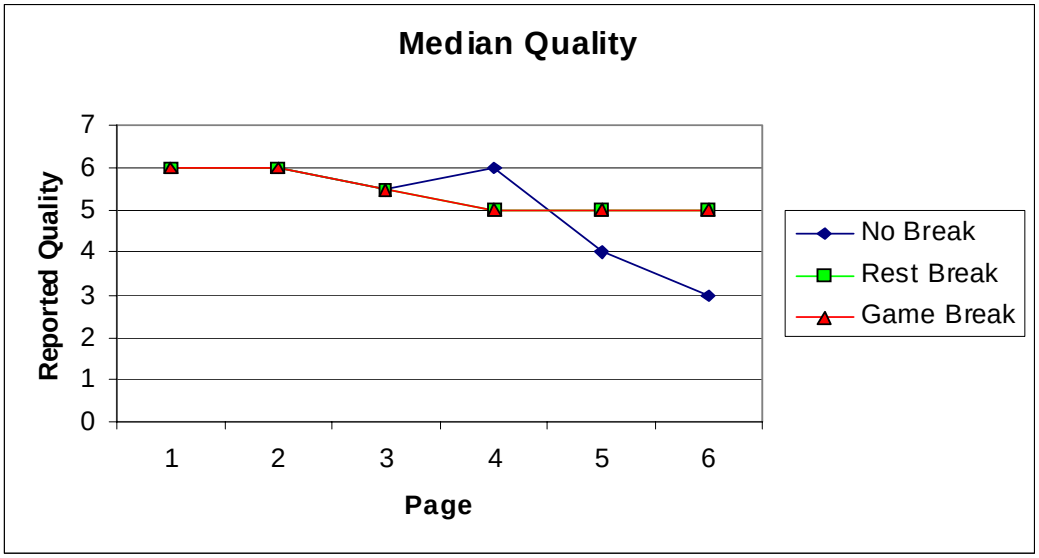


Figure 5.6: Pilot Study Quality per Page

Figure 5.6 plots the median quality response from each subject made on the y axis against the page number on the x axis. Each data set in the figure corresponds to a unique subject in the pilot study. The question asked was "How well do you think you copied the document, where a zero is many errors and ten represents a perfect copy?" In the groups that received breaks of any kind, quality remained steady throughout the entire process. Group A (no break) on the other hand, suffered a sharp drop in perceived quality after page four.

One of the most noticeable elements of the quality assessment was that the no break group differs from the other groups until page four. At page 4 only the no break group differs from the others two, first increasing in their perceived quality and then sharply decreasing. Though the real performance of these subjects was actually steady or improving over this time span, they may have been on the verge of entering their decay period even if it was not yet impacting their real performance. The two groups that got to take breaks were feeling more refreshed and maintained a fairly even assessment of their own quality.

5.4 Analysis

Overall, the data provided by the subject study does not provide support for the initial hypothesis. Looking at the raw productivity data (Fig. 5.3) there was no real difference between subjects that received a break and those that did not. The conclusion could be that rest breaks, with games or otherwise, do not have any impact on productivity. To make such a claim would be highly dubious since prior work in the field of productivity research has demonstrated that rest breaks (when correctly applied) enhance performance and productivity in a meaningful way[18,19]. There are a few behaviors that could account for the discrepancy between our results and the previous work on breaks and productivity, each of which is examined here in more detail.

Although transcription was chosen because it was thought that most people would be familiar with typing and using a computer, especially considering that our population was mostly college students at a technical school, it is possible that over the course of the study subjects were learning and improving at the task, offsetting any performance degradation due to exhaustion with performance improvement due to learning. While this seems unlikely, it is not impossible that subjects simply became accustomed to the writing style and diction of the article which would serve as a sort of learning and possibly improve performance. This is separate from the build up period pictured in Figure 1.1, but rather represents actual long term improvement at the task of transcription.

Related to this is the possibility is that the initial startup period resulting from taking a break is longer than anticipated, meaning that the performance degradation never took place because the subjects were not working for long enough. Other studies have

shown that performance degradation is measurable within an hour of starting to work[20]. Most of our subjects took roughly six minutes for each entry, plus one minute per entry to fill out the qualitative responses. Over six entries this adds up to only 42 minutes of work. It is possible that a clearer trend would have become evident if there were one or two more entries to transcribe in the study, but given our drop-out rates (Fig 5.1) there would have been very few participants left at that point.

A final possibility is that the ratio of work time to break time was too high. The ratio was based on the methods of *The Effect of Work-Rest Schedules*. We had a work time of 15 minutes followed by a 5 minute break followed by another 15 minutes of work. The previous study had 15 minutes of work followed by a 1-2 minute break. The use of this study as a benchmark might be inappropriate for our study.

Chapter 6: Conclusions

Playing computer games is sometimes perceived to have negative social and mental consequences. Because of this many workplaces are restricting worker access to Web sites that host computer games. However, many studies now show that playing computer games can in fact have positive side effects, particularly in the area of cognitive health. Companies that produce computer games have missed opportunities because of this bad reputation, and would like hard scientific data on the subject of games and productivity. If the reputation proves true, computer game companies could try to change their product or market appropriately. If instead, the reputation is not supported by data, then video game companies can use this information to promote broader acceptance of computer games in the workplace.

We designed and executed two studies based on other productivity research and applied to the unique constraints of playing computer games during work breaks. The pilot study examined the task of transcription. The goal was to learn how the productivity of transcribers changes over time as they continue to work without a break. This was successful; transcription appeared to create the necessary performance degradation within the space of an hour if no break was given. However, given our schedule of work/rest, an hour might not have been long enough to encompass the larger work/rest cycle proposed in the hypothesis.

The user study examined the effect a of computer game break on productivity using transcription as the work task. 42 Subjects transcribed an article, and each subject took a break to play a game, took a regular break, or took no break after the halfway point. Comparing the number of errors made by members of each group, the relative

productivity levels over the course of the task were determined. After careful analysis, we determined that the data collected by the user study was inconclusive in assessing the impact of computer games on productivity. At face value, the results indicated that a video game break had the same effect as a regular break, but that any break had the same impact as no break at all. However, this result is incongruous with previous productivity studies, which universally found that taking rest breaks increase overall performance. It is suspected that as subjects reached their own point of productivity decay, they quit the study. This meant that the only data collected was gathered from fully productive subjects. This is supported by the fast drop in quality reported by the no break group after the fourth page, indicating that this subject group was about to enter the decay period and leave the study. This conclusion is supported by the high rate of subjects dropping out of the study without finishing. We believe we do have the foundation for a technique that can be used in the future to examine the problem under more controlled circumstances and yield data to support or refute the reputation of computer games impact on productivity.

Though our overall methodology appears to be sound, the fine details of the execution may account for the discrepancy between our results and prior knowledge. The task of transcription, the three experimental groups, and the population of subjects with roughly equal word processing experience are a solid base from which to conduct this form of study. Numerous other studies have shown that transcribers do have a downward trend in performance over time, this supports the use of transcription as the task. However, in those studies, every participant finished the study; subjects were did not quit if they became dissatisfied with the task. We believe that our study did not show the

downward trend in productivity simply because as our subjects began to decline in performance, they stopped participating in the study. We speculate that had these subjects continued with the study, their downward trends would have pulled downward on the flat and/or upward trends that were being created by the subjects who had not yet reached their productivity decay point. Instead, we only see the flat and upward trends of those who have not yet reached their maximum tolerance.

Chapter 7: Future Work

Our work has provided an understanding of how to study the issue of how computer games affect worker productivity levels. With more time and more resources, future work could determine with finality the veracity of our hypothesis.

Future studies using our methodology should give subjects a better idea of how much work remains in their transcription task. We chose to not tell subjects how much work they had left to do; we feared that subjects who knew they were nearly done would perform differently than subjects who did not. We wanted to create conditions like a workplace, where work does not end after the first hour so chose not to inform subjects on their progress during the study. In retrospect, information about how many entries remain could reduce subject drop out rates. While many subjects will quit regardless of this change a progress bar might convince subjects who would otherwise have quit to continue onward, seeing how close they were to finishing and becoming eligible for the incentive.

Any future study constructed along the same lines as ours should be constructed in such a way that all the data considered is for complete passes through the study. Three tactics come to mind that will satisfy this constraint. The first option is to remove from the study any subjects who do quit partway through. This requires a much larger sample population than we were able to obtain, because the drop-out rate will cull out a large percentage of the subject pool. Also, if our hypothesis is correct, and the subjects who choose to continue do so because they have not yet begun to lose effectiveness. At that point, the only data being considered comes from people who have not yet begun to lose productivity as of the end of the study.

A more reliable solution to the problem of subjects dropping out is to not allow subjects to quit. This will require more draconian measures and/or greater resources to encourage participation, while trying to maintain the integrity of the data. In order to attract subjects for such a study, the incentive, monetary or otherwise, for each subject to participate in the study will have to be significant. Also, enforcing a policy of 'no quitting' will likely require that the study be taken off the Internet and into a more controlled setting, such as a specific computer lab at a scheduled time. This will make the study less accessible to some subjects who would otherwise be able to participate from home or work at their convenience.

Although these two alternatives were outside the limits of our time and resources, a company could implement either plan using its own employees as a testing base; the study would be a normal work break. A company also might have the financial resources to pay subjects for their time, which will most likely result in both a reduced drop out rate and a larger sample size.

Appendix

A: Pilot Study Pages

1

In the very olden time there lived a semi-barbaric king, whose ideas, though somewhat polished and sharpened by the progressiveness of distant Latin neighbors, were still large, florid, and untrammeled, as became the half of him which was barbaric. He was a man of exuberant fancy, and, withal, of an authority so irresistible that, at his will, he turned his varied fancies into facts. He was greatly given to self-communing, and, when he and himself agreed upon anything, the thing was done. When every member of his domestic and political systems moved smoothly in its appointed course, his nature was bland and genial; but, whenever there was a little hitch, and some of his orbs got out of their orbits, he was blander and more genial still, for nothing pleased him so much as to make the crooked straight and crush down uneven places.

Among the borrowed notions by which his barbarism had become semified was that of the public arena, in which, by exhibitions of manly and beastly valor, the minds of his subjects were refined and cultured.

But even here the exuberant and barbaric fancy asserted itself The arena of the king was built, not to give the people an opportunity of hearing the rhapsodies of dying gladiators, nor to enable them to view the inevitable conclusion of a conflict between religious opinions and hungry jaws, but for purposes far better adapted to widen and develop the mental energies of the people. This vast amphitheater, with its encircling galleries, its mysterious vaults, and its unseen passages, was an agent of poetic justice, in which crime was punished, or virtue rewarded, by the decrees of an impartial and incorruptible chance.

When a subject was accused of a crime of sufficient importance to interest the king, public notice was given that on an appointed day the fate of the accused person would be decided in the king's arena, a structure which well deserved its name, for, although its form and plan were borrowed from afar, its purpose emanated solely from the brain of this man, who, every barleycorn a king, knew no tradition to which he owed more allegiance than pleased his fancy, and who ingrafted on every adopted form of human thought and action the rich growth of his barbaric idealism.

When all the people had assembled in the galleries, and the king, surrounded by his court, sat high up on his throne of royal state on one side of the arena, he gave a signal, a door beneath him opened, and the accused subject stepped out into the amphitheater. Directly opposite him, on the other side of the inclosed space, were two doors, exactly alike and side by side. It was the duty and the privilege of the person on trial to walk directly to these doors and open one of them. He could open either door he pleased; he was subject to no guidance or influence but that of the aforementioned impartial and incorruptible chance. If he opened the one, there came out of it a hungry tiger, the fiercest and most

cruel that could be procured, which immediately sprang upon him and tore him to pieces as a punishment for his guilt. The moment that the case of the criminal was thus decided, doleful iron bells were clanged, great wails went up from the hired mourners posted on the outer rim of *the arena, and the vast audience, with bowed heads and downcast hearts, wended slowly their homeward way, mourning greatly that one so young and fair, or so old and respected, should have merited so dire a fate.

2

On January 3, 1920, Boston Red Sox owner Harry Frazee sold Babe Ruth to the New York Yankees for $125,000. The entire Yankees-Red Sox dynamic--the Yankees' eternal supremacy, the Red Sox's eternal heartbreak--can be traced back to this fateful moment. Before 1920, the Red Sox had won five World Series, the Yankees none. Since then? The Yankees 26, the Red Sox none. Some would even call it a curse.

As thoroughly and heartbreakingly as the Yankees have dominated Boston, the Red Sox Nation is forever sustained by the notion that this, finally, is the year and oh, how sweet it's going to be. "If you can win a World Series," former Red Sox catcher Rich Gedman once said, "you will he immortal around here.

And the Yankees? Just like Road Runner out-foxing Wife E. Coyote, they are sustained by the pure, unadulterated joy of foiling the Sox, over and over. Yogi Berra once put it, with uncharacteristically unbungled eloquence. "We've been playing these guys for 80 years. They're never going to beat us." Beep! Beep!

The Red Sox don't buy that. Can't buy that. So, after finishing second in the A.L. East behind the Yankees for the past six seasons, they went hellbent for a pennant and World Series title last offseason, and have assembled a team equipped to do so, if things break right. Like they ever do, But you never know.

Indeed, it seems clear, as the Red Sox and Yankees had their first meeting of the 2004 season in mid-April at Fenway Park, with Boston winning three of the four games indicating the rivalry his never been more heated, more robust and more competitive.

"If you go and look at a football season, where one game seems to be a turning point, it's a good analogy," Yankees manager Joe Torre said. "One game stands out. There's 19 seasons here, when you play these guys.

"Every single game is ... I hesitate to say warlike, because you certainly don't want to get into the feeling it's dangerous out hare. But I think the passion and need to win on both sides is something you won't experience anywhere else."

The Red Sox and Yankees played an unprecedented 26 times last year, including those seven in the post season. But now the rivalry extends even to spring training, where fans slept out overnight to get tickets to their "showdown" in Fort Myers, Florida last March.

Tickets were being scalped for $200, and commemorative pins were sold to mark the occasion.

Now the hype is nearly nonstop. Five hooks pertaining to the rivalry are scheduled for publication, along with two film documentaries and an off-Broadway play.

The rivalry has seeped into the popular culture, from references in Adam Sandier and Matt Damon movies to "Catch Me if You Can," in which Frank Abagnale's father asks his son, "Do you know why the Yankees always win? Because the other team's too busy staring at the pinstripes."

3

"We lit a fire, offered some of the cheeses in sacrifice, ate others of them, and then sat waiting till the Cyclops should come in with his sheep. When he came, he brought in with him a huge load of dry firewood to light the fire for his supper, and this he flung with such a noise on to the floor of his cave that we hid ourselves for fear at the far end of the cavern. Meanwhile he drove all the ewes inside, as well as the she-goats that he was going to milk, leaving the males, both rams and he-goats, outside in the yards. Then he rolled a huge stone to the mouth of the cave- so huge that two and twenty strong four-wheeled waggons would not be enough to draw it from its place against the doorway. When he had so done he sat down and milked his ewes and goats, all in due course, and then let each of them have her own young. He curdled half the milk and set it aside in wicker strainers, but the other half he poured into bowls that he might drink it for his supper. When he had got through with all his work, he lit the fire, and then caught sight of us, whereon he said:

"'Strangers, who are you? Where do sail from? Are you traders, or do you sail the as rovers, with your hands against every man, and every man's hand against you?'

"We were frightened out of our senses by his loud voice and monstrous form, but I managed to say, 'We are Achaeans on our way home from Troy, but by the will of Jove, and stress of weather, we have been driven far out of our course. We are the people of Agamemnon, son of Atreus, who has won infinite renown throughout the whole world, by sacking so great a city and killing so many people. We therefore humbly pray you to show us some hospitality, and otherwise make us such presents as visitors may reasonably expect. May your excellency fear the wrath of heaven, for we are your suppliants, and Jove takes all respectable travellers under his protection, for he is the avenger of all suppliants and foreigners in distress.'

"To this he gave me but a pitiless answer, 'Stranger,' said he, 'you are a fool, or else you know nothing of this country. Talk to me, indeed, about fearing the gods or shunning their anger? We Cyclopes do not care about Jove or any of your blessed gods, for we are ever so much stronger than they. I shall not spare either yourself or your companions out of any regard for Jove, unless I am in the humour for doing so. And now tell me where

you made your ship fast when you came on shore. Was it round the point, or is she lying straight off the land?'

"He said this to draw me out, but I was too cunning to be caught in that way, so I answered with a lie; 'Neptune,' said I, 'sent my ship on to the rocks at the far end of your country, and wrecked it. We were driven on to them from the open sea, but I and those who are with me escaped the jaws of death.'

4

The railway underwent two changes of name during this time period. In 1972, the railway's name was changed to the British Columbia Railway (BCR). In 1984, the BCR was restructured. Under the new organization, BC Rail Ltd. was formed, owned jointly by the British Columbia Railway Company (BCRC) and by a BCRC subsidiary, BCR Properties Ltd. The rail operations became known as BC Rail.

In 1973, the British Columbia government acquired and restored an ex-Canadian Pacific Railway 4-6-4 steam locomotive of the type known as "Royal Hudsons", a name that King George VI permitted the class to be called after the Canadian Pacific Railway used one on the royal train in 1939. The locomotive that the government acquired, numbered 2860, was built in 1940 and was the first one built as a Royal Hudson. The government then leased it to the British Columbia Railway, which started excursion service with the locomotive between North Vancouver and Squamish on June 20, 1974. The train ran between June and September on Wednesdays through Sundays.

In the 1960s, a new line had been projected to run northwest from Fort St. James to Dease Lake, 412 miles (623 km) away. On October 15, 1973, the first 125 miles (200 km) of the extension to Lovell were opened. The cost of the line was significantly greater than what was estimated, however. Contractors working on the remainder of the line alleged that the railway had misled them regarding the amount of work required so that it could obtain low bids, and took the railway to court.

The Dease Lake line was starting to appear increasingly uneconomical. There was a world decline in the demand for asbestos and copper, two main commodities that would be hauled over the line. As well, the Cassiar Highway that already served Dease Lake had recently been upgraded. Combined with the increasing construction costs, the Dease Lake line could no longer be justified. Construction stopped on April 5, 1977. Track had been laid to Jackson, 263 miles (423 km) past Fort St. James, and clearing and grading were in progress on the rest of the extension. It had cost $168 million to that point, well over twice the initial estimate.

The management and operation of the railway had been called into question, and on February 7, 1977, the provincial government appointed a Royal Commission, the McKenzie Royal Commission, to investigate the railway. Its recommendations were released on August 25, 1978. It recommended that construction not continue on the 149 miles (239 km) of roadbed between Dease Lake and the current end of track, and that

trains be terminated at Driftwood, 20 miles (33 km) past Lovell. The rest of the track would be left in place but not used. In 1983, after logging operations ceased at Driftwood and traffic declined sharply, the Dease Lake line was closed. However, it was reopened in 1991 and, as of 2005, extends to a point called Chipmunk, still over 175 miles (280 km) south of Dease Lake. Many of the Commission's other recommendations, including the abandonment of the Fort Nelson line, and discontinuation of uneconomic operations such as passenger services, were not followed. In the early 1980s the railway built a new line and acquired another. The Tumbler Ridge Subdivision, an 82 mile (132 km) electrified branch line, opened in 1983 to the Quintette and Wolverine mines, two coal mines northeast of Prince George that produced coal for Japan.

5

Expedient sheltering involves the use of common materials to enhance the safety of a room inside a building against the impacts of a chemical plume. The central premise behind taping and sealing with duct tape and plastic is to reduce airflow into a room. Vapors penetrate into a room through cracks and openings in the walls, floors and ceilings, around doors and windows, and through openings for ducts, light fixtures, fans, pipes, electrical outlets, chimneys, door handles, and locks. The goal of taping and sealing is to significantly reduce infiltration at these points.

Expedient sheltering was suggested by NATO (1983) using the term “ad-hoc shelter” to protect civilian populations from chemical warfare agent exposure. The concept was to use plastic sheeting to seal off a room by fashioning a simple airlock at the entrance to the room and sealing off doors, windows or louvered vents. The NATO guidelines also stressed the need for rapid exit from the ad-hoc shelter once the plume had passed to avoid further exposure (NATO 1983, p. 143).

This strategy was further developed by the Israeli Civil Defense in the mid-1980s to protect the public against a chemical weapons attack (Yeshua 1990). The tape and seal strategy was in place when the Gulf War occurred in 1991 and received considerable media attention. The Israeli strategy was to have citizens prepare a “safe room” in their house or apartment with the use of weatherization techniques to permanently reduce infiltration. Citizens were also instructed to take expedient measures, such as sealing doors and windows with plastic sheets, in the event of a chemical weapons attack. The use of plastic over a window was developed to reduce air infiltration and to provide a vapor barrier in the event of glass breakage from bomb explosions. A modification of the Israeli strategy was proposed for use in CSEPP (Sorensen 1988; Rogers et al. 1990).

Although vapors, aerosols, and liquids cannot permeate glass windows or door panes, the amount of possible air filtration through the seals of the panes into frames could be significant, especially if frames are wood or other substance subject to expansion and contraction. To adequately seal the frames with tape could be difficult or impractical. For this reason, it has been suggested that pieces of heavy plastic sheeting larger than the window be used to cover the entire window, including the inside framing, and sealed in place with duct or other appropriate adhesive tape applied to the surrounding wall.

Another possible strategy would be to use shrink-wrap plastic often used in weatherization efforts in older houses. Shrink-wrap commonly comes in a 6 mil (0.006-in.) thickness and is adhered around the frame with double-faced tape and then heated with a hair dryer to achieve a tight fit. This would likely be more expensive than plastic sheeting and would require greater time and effort to install. Because double-faced tape has not been challenged with chemical warfare agents, another option is to use duct tape to adhere shrink-wrap to the walls. Currently, we do not recommend using shrink-wrap plastics because of the lack of information on its suitability and performance.

Duct tape and plastic sheeting (polyethylene) were chosen because of their ability to effectively reduce infiltration and for their resistance to permeation from chemical warfare agents.

B: User Study Entries

1

In a stinging rebuke to President Bush's war strategy, the Senate called on the White House yesterday to provide regular updates on the conditions for withdrawal in Iraq and voted to allow some terrorists convicted by military tribunals at US detention camps in Guantanamo Bay, Cuba, to appeal their military verdicts in civilian courts.

Both the Iraq proposal and the Guantanamo provision, which allows a terror suspect to appeal if a military tribunal sentences him to death or more than 10 years in prison, passed by overwhelming margins. The rare bipartisan vote signals that Republicans and Democrats will no longer give the White House carte blanche in applying its wartime authority.

Sponsored by Senator John Warner of Virginia, the influential Republican chairman of the Armed Services Committee, the Iraq amendment stopped short of demanding a timetable for withdrawal, as Democrats had wanted. But the amendment, cosponsored by Senate majority leader Bill Frist, a Tennessee Republican, declares that "US military forces should not stay in Iraq any longer than required and the people of Iraq should so be advised."

2

Graham said the amendment "allows every detainee under our control to have a day in court" and can help prevent inmates and their lawyers from clogging the courts with what he deemed frivolous petitions the anticipated byproduct of a Supreme Court decision giving them the right to appeal their detention.

Rights groups applauded lawmakers yesterday for taking on the issue, but said Graham's measure does more harm than good because it effectively denies due process to other terror suspects who haven't been tried or whose sentences don't fit the standard.

It also supersedes the high court's ruling and bars detainees from petitioning civilian courts about their original incarceration, critics charged. The Bush administration has reserved the right to hold war detainees indefinitely without trial and has fought to keep terror suspects out of the US justice system.

"Instead of fixing the military commission problems, they narrow the scope of review," said Elise Massimino, Washington director of Human Rights First, a liberal nonprofit organization. "The real problem is that the administration claims it can hold people forever, without charging them and without due process of law."

3

Indeed, public opinion polls indicate that support for Bush's handing of the war has plummeted in recent months as US military deaths have climbed past 2,000 and increasingly sophisticated insurgent attacks seem to continue unabated. While Democrats have been more critical of the war, the White House's GOP allies in Congress are now showing deepening signs of angst.

Those concerns seem to extend to the president's handling of his overall wartime authority, including how to mete out justice to suspected members of Al Qaeda and other terrorist groups.

Immediately after approving Warner's amendment, senators approved, by an 84-to-14 tally, a second measure sponsored by Senator Lindsey O. Graham, Republican of South Carolina and a military judge in the Air Force Reserves.

Graham's proposal would grant convicted terrorists limited access to civilian courts to appeal their convictions, a way to provide more oversight of the judicial process for detainees, which has been the focus of international condemnation. Kennedy and Leahy were the lone New England no votes.

4

"People who are willing to risk their lives [in the military] need to know the truth," he said. "They need to understand that they are there based on decisions that were made in good faith by responsible people."

Still, Warner said his amendment, approved 79 to 19, sends a "very powerful statement" to the Iraqis that American forces have done what they can and Iraqis themselves must now stand up against the insurgency. He said he was "very grateful" for the widespread support. Of the 12 senators from New England, Democrats Edward M. Kennedy and John F. Kerry of Massachusetts and Democrat Patrick J. Leahy of Vermont were the only ones to vote against the amendment. Staff members said the senators did not believe the amendment was strong enough.

Others said the vote underscored growing public demands for a clear exit strategy.

"After two-and-a-half years of insurgency warfare in Iraq, it is a stunning indictment of the Bush administration that this Senate has to ask for a plan," said Senator Jack Reed, a Democrat from Rhode Island. "And we're asking on behalf of the American people. Because there is disquiet with Iraq."

5 (note that due to an error, this page was never displayed, subjects saw 1 again instead)

The amendment also urges Bush to press Iraq's political and religious groups to work out their differences "essential for defeating the insurgency in Iraq," according to the amendment.

Secretary of Defense Donald H. Rumsfeld said yesterday that he understands lawmakers' desire for more information, but he warned against rushing US troops home before the Iraqis are ready to fight.

"Timing of the handover of responsibility to Iraqis depends on conditions on the ground," Rumsfeld said. "And already some responsibilities are being assumed by the Iraqi security forces. We must be careful not to give terrorists the false hope that if they can simply hold on long enough, that they can outlast us."

Repeating a recent White House theme, Rumsfeld lashed out at Democrats who are accusing Bush of exaggerating intelligence and misleading the country about the rationale for going to war. Rumsfeld quoted prominent Democrats and Republicans who saw the same intelligence as the White House and concurred that Iraq's weapons of mass destruction which have yet to be found were an intolerable threat.

Warner called the vote a "strong bipartisan message to the world" that it was time for Iraqis to take control.

"The coalition forces, most particularly the United States and Great Britain, have done their job," Warner said after the vote. "And now we expect in return that [the Iraqis] take charge of their nation and run it and form a democracy and prevent any vestige of a civil war from taking place." t

The measures won't reach Bush's desk unless the House also adopts them and some think that's a long shot.

"This probably won't pass the House, but even so it sends a signal that Senate Republicans are becoming more demanding," said James Phillips, a Middle East analyst at the conservative Heritage Foundation. "Republican senators are more actively second-guessing the administration."

Added to a military spending bill, Warner's amendment calls for Bush to give public progress reports on the war every three months, and sets 2006 as the deadline for the United States to give Iraqis responsibility for their own security. "It is our intent to create the conditions in that year for the phased redeployment of the United States forces from Iraq," it stated.

C: Advertisement / Flyer

HELP with IQP!

WIN a PRIZE!

PLAY a GAME!

Hello,

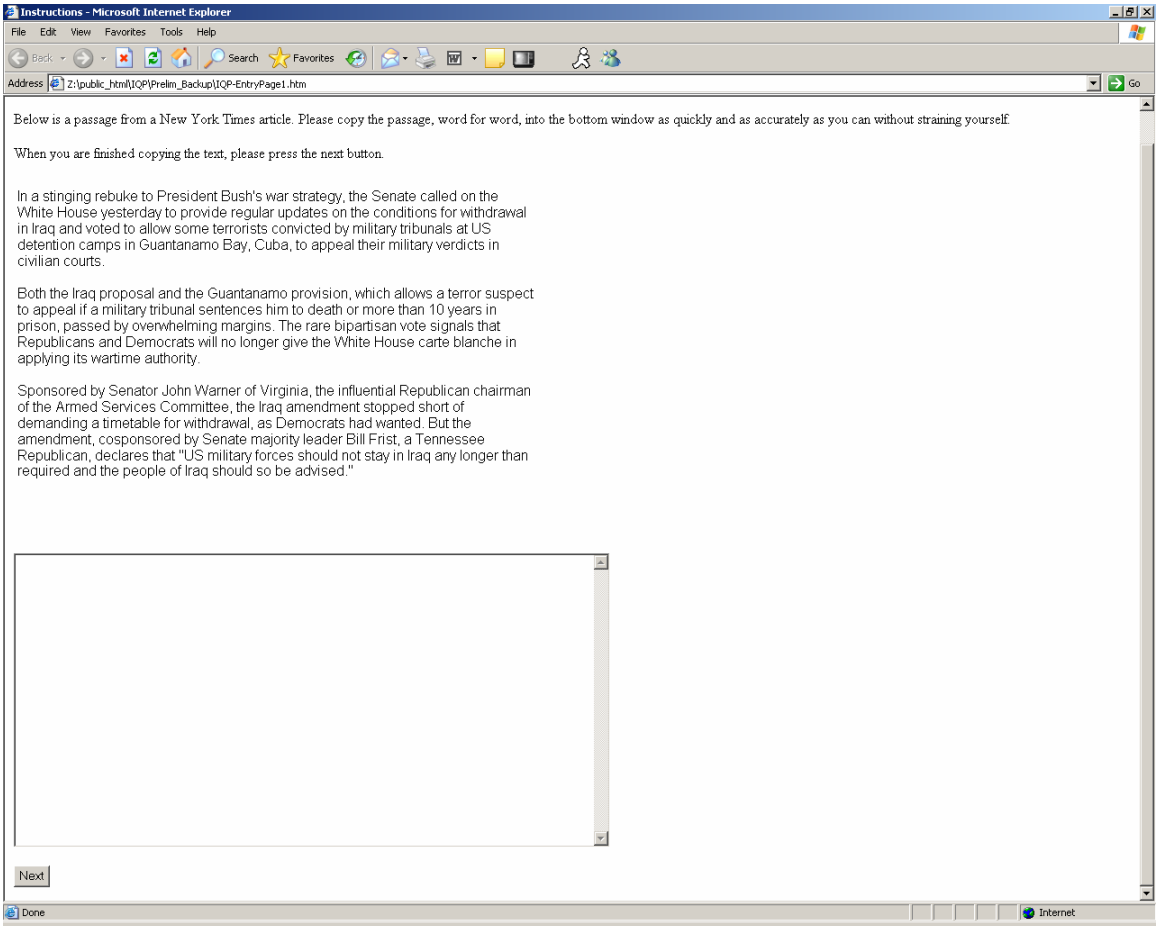
We need volunteers for an IQP study on video games and work productivity. Participants will do simple task and report on their progress. The study takes 30 minutes to 1 hour to complete and all participants are eligible to be entered into a drawing for a **$25 Best Buy** gift certificate. If you are interested in participating, please follow the link below.

http://www.wpi.edu/~mikea/IQP/Prelim_Files/startstudy.htm

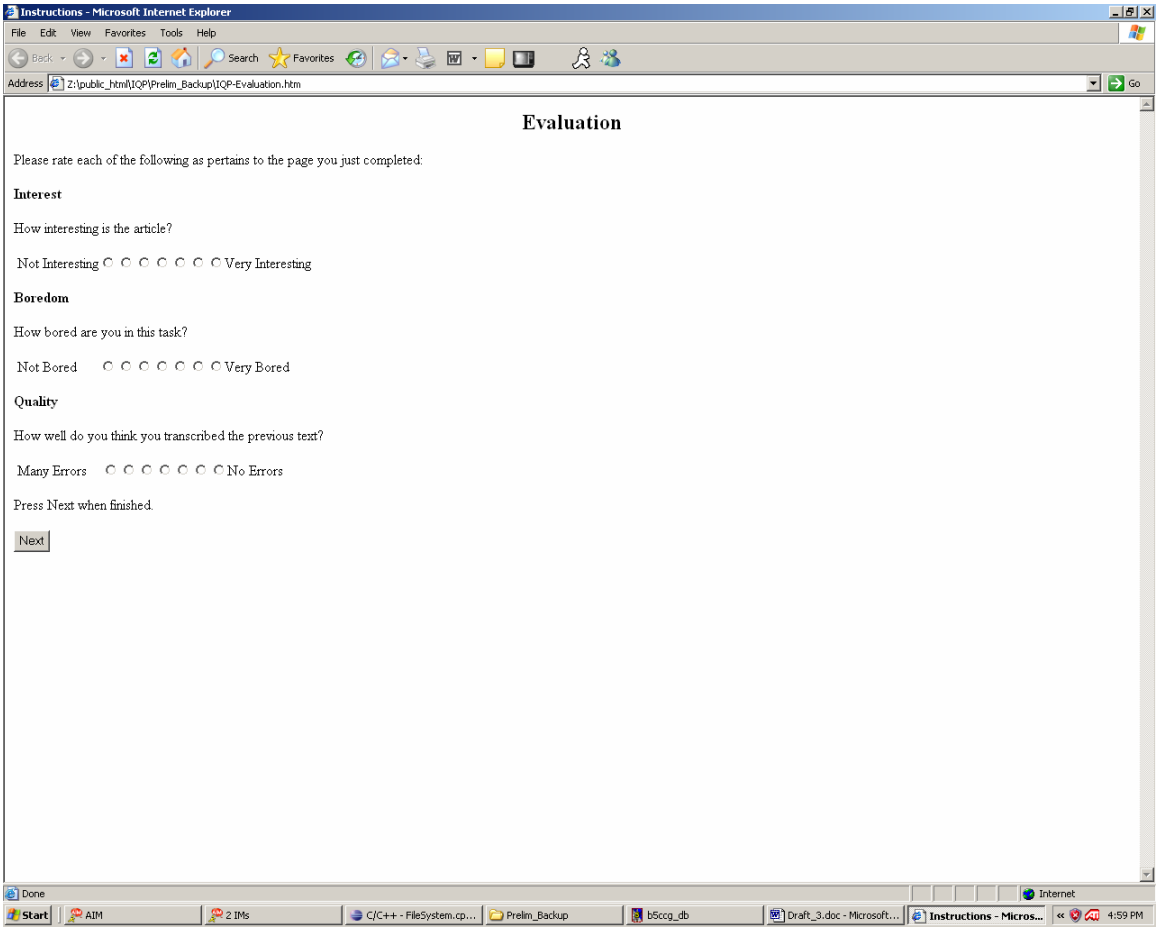
Jeremiah Chaplin (jchaplin@wpi.edu)
Michael Anastasia (mikea@wpi.edu)
Prof. Mark Claypool (claypool@cs.wpi.edu)

D: Web Interface

Interface for transcription



Interface for qualitative data entry



E: Additional Data gathered in User Study

---

[1] Gaspar, Lesley, *Mouse Trap*, Gale Group, Inc., 2006

[2] Phoon, Annie, *Memory Message: Review Games That Enhance Retention*, Technical & Skills Training 8, no. 1 (January 1997), Gale Group Inc., 2006

[3] Gaspar, Lesley, *Mouse Trap*, Gale Group, Inc., 2006

[4] Ben Sawyer, *PopCap(TM) and Games For Health Project Announce Joint Research Initiative on Cognitive Health Benefits of Digital Gameplay*, Games for Health, http://www.gamesforhealth.org, 2006

[5] Miller, Greg, Computer Game Sharpens Aging Minds, Society for Neuroscience Meeting News of the Week,Science 25 November 2005, Vol. 310. no. 5752, p. 1261, DOI: 10.1126/science.310.5752.1261a

[6] Branham Group, *Developer Producitivity Study*, Branham Group Inc.100 Constellation Crescent, Suite 915 Ottawa, ON K2G 7E6, 2005

[7] M.-H. Liao, C. G. Drury , *Posture, discomfort and performance in a VDT task*, Volume 43, Number 3, Taylor & Francis, 2000

[8] Rana, Balci and Fereydoun, Aghazadeh. *Effects of Exercise Breaks on performance, muscular load, and perceived discomfort in data entry and cognitive tasks.* Department of Industrial and Manufacturing Systems Engineering, Louisiana University

[9] Rana, Balci and Fereydoun, Aghazadeh. *Effects of Exercise Breaks on performance, muscular load, and perceived discomfort in data entry and cognitive tasks.* Department of Industrial and Manufacturing Systems Engineering, Louisiana University

[10] , Tucker P, Folkard S, Macdonald I. *Rest breaks and accident risk,* Department of Psychology, University of Wales, SA2 8PP, Swansea, UK. , 2003

[11] ,Rossa, RR, *Extended Workshifts and Excessive Fatigue*, National Institute for Occupational Safety and Health, Division of Biomedical and Behavioral Science, Cincinnati, Ohio, USA, 1995

[12] Henning, RA; Kissel, GV; Maynard, DC, *Compensatory rest breaks for VDT operators*
International Journal of Industrial Ergonomics [INT. J. IND. ERGONOMICS]. Vol. 14, no. 3, pp. 243-249. 1994.

[13] Griffith JL, Voloschin P, Gibb GD, Bailey JR, *Differences in Eye-Hand Motor Coordination of Video-Game Users and Non-Users*, Percept Mot Skills (0031-5125); Volume: 57; Issue: 1, 1983

[14] Ricardo Rosasa,*, Miguel Nussbaumb, et al., *Beyond Nintendo: design and assessment of educational computer games for first and second grade students,*aSchool of Psychology, Pontificia Universidad Cato´lica de Chile, Santiago, Chil and School of Engineering, Pontificia Universidad Cato´lica de Chile, Santiago, Chile, 29 July 2002

[15] Gaspar, Lesley, *Mouse Trap*, Gale Group, Inc., 2006

[16] zyphrus, *TubeBlazer*, http://sourceforge.net/projects/tubeblazer

[17] Brian Postma, *OpenGL Pacman,* http://www.brianpostma.com/pc.html

[18] Rana, Balci and Fereydoun, Aghazadeh. *Effects of Exercise Breaks on performance, muscular load, and perceived discomfort in data entry and cognitive tasks.* Department of Industrial and Manufacturing Systems Engineering, Louisiana University

[19] ,Rossa, RR, *Extended Workshifts and Excessive Fatigue*, National Institute for Occupational Safety and Health, Division of Biomedical and Behavioral Science, Cincinnati, Ohio, USA, 1995

[20] Rana, Balci and Fereydoun, Aghazadeh. *Effects of Exercise Breaks on performance, muscular load, and perceived discomfort in data entry and cognitive tasks.* Department of Industrial and Manufacturing Systems Engineering, Louisiana University