Dealing with Distractions: Analyzing and Designing for Task Switching at Work

Elisabeth Sylvan
B.S. in Psychology, Carnegie Mellon University 1996

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning, in partial fulfillment of the requirements for the degree of Master of Science in Media Arts and Sciences at the Massachusetts Institute of Technology June 2003

Massachusetts Institute of Technology 2003. All rights reserved.
Dealing with Distractions: Analyzing and Designing for Task Switching at Work

Elisabeth Sylvan

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning, in partial fulfillment of the requirements for the degree of Master of Science in Media Arts.

Abstract

Increasingly, distraction is just a click away: computers offer you the temptation to switch from your primary task. The current work describes four investigations into how well people manage switching between tasks and between multiple computer applications: a controlled experiment on task-switching, a survey on email behavior and efficient, a survey on how different computer applications are used for work and outside of work, and an analysis of computer users real-world switching behavior. Based on the findings, two outcomes are presented: 1.) a new software tool that visualizes workflow to encourage users to reflect upon their work habits and their task switching and 2.) implications for the design of monitoring software, tools for self-reflection, and push or disruptive technologies such as email systems.

Keywords: HCI, task switching, distractions, interruptions, productivity, work habits, push technology, disruptive technology.

Thesis Advisor: Dan Ariely
Dealing with Distractions: Analyzing and Designing for Task Switching at Work

Elisabeth Sylvan
Master of Science Thesis
June 2003

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning, in partial fulfillment of the requirements for the degree of Master of Science in Media Arts and Sciences at the Massachusetts Institute of Technology 2003. All rights reserved.

The following people served as readers for this thesis:

Advisor: 
Dan Ariely
Assocate Professor of Behavioral Economics and Media Arts and Sciences MIT Sloan School of Management and the Program in Media Arts and Sciences

Reader: 
Rosalind Picard
Associate Professor of Media Arts and Sciences Program in Media Arts and Sciences

Reader: 
Bakhtiar Mikhak
Research Scientist MIT Media Lab
Acknowledgements

This thesis would not have been possible without the support of many people both at the Lab and outside.

My advisor, Professor Dan Ariely, surprises and delights me with his unusual worldview and distinctive research. He regularly asks me whether I'm having fun, an unusual question from an advisor. He has challenged my idea of what my work is and should be, and though I've pushed back, he's made me a more academically minded, and I believe, better Media Lab student.

Beginning with my first interactions with him, my reader, Dr. Bakhtiar Mikhak has challenged me to reevaluate my beliefs, both personal and scholarly. His broad and deep knowledge of how our world is constructed and how we construct it, his ability to apply what he knows in the most caring and careful way, and his quiet, intense passion are all an inspiration to me. Like many students at the lab, I feel very lucky to know him.

Professor Rosalind Picard has been a wonderful support long before she graciously agreed, despite her full schedule, to read my thesis. As a regular sounding board for my research, she's given much appreciated, thoughtful advice on everything from ethical research practice to how to achieve the delicate balance between theory and implementation.

Linda Peterson has been an extraordinary resource throughout the program, a unique and critical source of support. The lab is lucky to have her.

My UROP, Kaska's cheerful smile and diligent work has brightened my days for the years she's worked with me and I know that she will be a success at whatever she sets her mind to accomplish.

Thanks to the many wonderful students and other coconspirators who have provided support, advice and laughs along the way including Mike A. and Mike N., Jennifer S. and Jennifer B. Michelle, Cory, Dona and Joanie.

Thanks to Mom, Dad, and Jeremy for their love and support through the years.

And to my love and my husband, Misha, without whom none of this is possible.
## CONTENTS

1 INTRODUCTION ........................................................................7

2 BACKGROUND LITERATURE ....................................................... 15
   2.1 Attention, cognitive load and task switching ..............................................15
   2.2 Disruptive software applications .............................................................16
   2.3 Time discounting and impatience ............................................................18

3 TWO STUDIES ON TASK SWITCHING ............................................. 18
   3.1 Experiment on Task Switching ................................................................19
      3.1.1 Methods...................................................................................... 19
      3.1.2 Results ...................................................................................... 22
      3.1.3 Discussion.................................................................................... 25
      3.1.4 Conclusion................................................................................... 26
   3.2 How Efficiently Power Users Manage their Email .........................................27
      3.2.1 Methods...................................................................................... 28
      3.2.2 Results ....................................................................................... 29
      3.2.3 Discussion.................................................................................... 40
      3.2.4 Conclusions.................................................................................. 41

4 SYSTEM DEVELOPMENT ............................................................ 42
   4.1 Investigations into Work Habits .............................................................42
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1.1</td>
<td>Some Existing Work</td>
<td>43</td>
</tr>
<tr>
<td>4.1.2</td>
<td>Survey on People’s Work Habits</td>
<td>44</td>
</tr>
<tr>
<td>4.1.3</td>
<td>Observing Task Switching in Computer Work</td>
<td>51</td>
</tr>
<tr>
<td>4.2</td>
<td>Design Outcomes and Implications of These Findings</td>
<td>59</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Monitoring Application Switching</td>
<td>59</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Thoughts on Monitoring Tools Used By Organizations and Individuals</td>
<td>62</td>
</tr>
<tr>
<td>4.2.3</td>
<td>One Prototype Based on these Findings</td>
<td>65</td>
</tr>
<tr>
<td>4.2.4</td>
<td>Specific Design Ideas for Systems that Support Self Aware Work Habits</td>
<td>68</td>
</tr>
<tr>
<td>4.2.5</td>
<td>General Implications for User Interface Designers</td>
<td>71</td>
</tr>
<tr>
<td>5</td>
<td>CONCLUSIONS</td>
<td>75</td>
</tr>
<tr>
<td>6</td>
<td>REFERENCES</td>
<td>78</td>
</tr>
<tr>
<td>7</td>
<td>APPENDIXES</td>
<td>82</td>
</tr>
<tr>
<td>7.1</td>
<td>Task-switching Study Instructions</td>
<td>82</td>
</tr>
<tr>
<td>7.2</td>
<td>Email Survey Questions</td>
<td>86</td>
</tr>
<tr>
<td>7.3</td>
<td>Work Habits Survey Questions</td>
<td>88</td>
</tr>
<tr>
<td>7.4</td>
<td>Excerpt of an Example WorkSniffer Log File</td>
<td>90</td>
</tr>
<tr>
<td>7.5</td>
<td>Essay Experiment Instructions</td>
<td>91</td>
</tr>
<tr>
<td>7.6</td>
<td>Work Logging Observation Instructions</td>
<td>92</td>
</tr>
</tbody>
</table>
1 Introduction

Yet we have all experienced time when, instead of being buffeted by anonymous forces, we do feel in control of our actions, masters of our own fate. On the rare occasions that it happens, we feel a sense of exhilaration, a deep sense of enjoyment that is long cherished and that becomes a landmark in memory for what life should be like. The best moments usually occur when a person’s body or mind is stretched to its limits in a voluntary effort to accomplish something difficult and worthwhile. Optimal experience is thus something we make happen.

—Mihaly Csikszentmihalyi

Now imagine how you work when you write with your computer. Perhaps for a while the words flow easily. Then you reach a difficult section. You aren’t sure how to communicate your message so you decide to google it. As you search, you wander farther from your primary goal of writing a paper. Finally you get back to work. You are writing slowly, but writing, nonetheless. Then your computer beeps to tell that you’ve received new mail. Since you are curious who wrote you, you switch to find you have received a boring memo about an upcoming meeting. When you return to your paper, you can’t remember exactly what idea you were constructing before.

Does this sound familiar?

Computers have radically changed how we can work. We write without pens. We chat without saying a word. Everything is now in one place: our tools, our toys, our communication mechanisms. We write our papers, build our software, do our taxes, play solitaire, write our mothers all with one device. Our tools
have changed so profoundly we may wonder—or perhaps not remember—how we used to get things done.

Our work process has changed with the tools. Email, instant messages, and other communication tools allow us to stay connected with friends, family and associates throughout our day. Calendar systems and project management tools help us to keep our busy lives organized by reminding us what we should be doing when.

But there is a drawback to the flexibility of our preferred tool. We are often interrupted by messages competing for our attention and have ample opportunity for us to diverge from our primary work with easy access to a variety of engaging activities such as web surfing, reading email, and game playing. We are interrupted at the wrong time by information of variable importance that is poorly matched to our current task.

What does this profound flexibility mean in term of how we work? What is that work process like? What needs improvement and what doesn’t? What can designers do to help busy people with divided attentions?

I propose that computers may be well suited for goals in which we switch quickly from one task to another. But they may be poorly suited for tasks that require focusing on one idea. Activities that require deep thought—such as
programming and writing – often require extended periods of uninterrupted focus. Developing our ideas may require keeping track of the big picture while working through the details. We may feel that we need a certain amount of mental, physical, and social space to allow our ideas to coalesce. Maintaining this space can be a battle and deep work requiring intense concentration may be particularly vulnerable to distractions. It is easy to switch to another task when the thinking gets hard. The sound of an email alert may cause us to lose attention. As we continue to work, repressing our thoughts about a new email surprise, our attention may be divided by the temptation of switching. If our thoughts start to stray and we begin thinking about poking around the web for inspiration, we may lose track of the mental structures we meant to build.

I suggest that when distraction is always a click away, we may not switch as efficiently as we would like. We may move from one task to another, barely noticing how much we switch and how switching impacts our work. We may not consider the cost of switching. If work isn’t fun but, say, email is, then we are tempted to read email and may not consider the problems of putting off work, just for a minute. Perhaps the immediate reward of enjoyable activities makes it persuasive. Maybe our attention moves away from our current task when we notice a message is waiting or we see our browser window open behind our programming environment. Or maybe we are just curious.
But switching may not be all bad. Even during the most intense projects, we are likely to have other important tasks that may require attention. When we need to take a break—particularly from work that is tedious or cognitively intense—we can attend to these other items. This break may help us to have a fresh perspective when we return to our work.

Sometimes the information that interrupts us is relevant or important, for instance, when interruptions are about the tasks that one we are currently undertaking. If, while programming, I receive an email from a coworker about fixing the bug that’s driving me crazy, I certainly appreciate it. Other times the juxtaposition of different subjects may lead us to invent or create new ideas. If I have writer’s block and don’t know what to say next and suddenly a consortium talk description arrives in my inbox and sparks my imagination, I am tickled.

Our work habits impact how much and what we accomplish during our days. They also may impact how we feel about our work and how deeply we are able to work. For a variety of reasons, understanding people’s work processes and supporting them through organizational and interface design is tricky. Every person may develop a system of working, particular to herself. People may have a range of personal styles. Similar styles may be more or less successful for different people. (For instance, younger people growing up with computers as tools intrinsic to their environment may thrive on the ease of switching and
have little patience for restrictive work environments.) Different types of work may require different work strategies. People may be unaware or unwilling to change their habits so that they may work more effectively and efficiently. Identifying which aspects of particulars styles are successful may be difficult.

This thesis addresses how people switch tasks. It asks what task switching looks like, whether and when people switch well, and how productivity is influenced by task switching. When is switching an effective means of taking a break or getting relevant information and when is it a distraction? A series of four studies ask these questions in a variety of ways. Based on the results of the investigations, two outcomes are presented: 1.) a software application that supports efficient task-switching through self reflection on one's work habits and 2.) some implications for organizations and user interface designers.

A few comments on the scope and direction of this work seem important to mention. A world exists beyond a person and her computer. Phones ring. Coworkers stop by. Errands need to be run and families need love and attention. Many researchers have studies how people manage distractions in their physical environment (cf. Gruen 1996) and this topic deserves revisiting. For simplicity and scale, this thesis focuses upon interruptions and distractions on the screen only.
Also this thesis uses the literature on attention, dual tasks and cognitive load, as a foundation and addresses a different question than these topics of study do. The issue at hand is when people choose to switch tasks and how this decision affects their productivity, not what they can attend to and what is the cost of switching. I am interested in how people distract themselves when faced with a potential interruption and how this work pattern affects their productivity. This is more an issue of decision-making than it is one of information processing.

Finally, I will make a distinction between an interruption and a distraction. For this document an interruption is an incoming bit of information different than the task at hand. A distraction is when a person switches from a primary task to a secondary one.

This thesis comprises four basic sections. This first section introduces the problem and its scope, explains the relevance of this problem, and describes the structure of the remaining document.

Section two describes the relevant background literature on how people work and how their habits, their environment, and their abilities all impact how they accomplish and how they feel. Recent work on how interruptions occur in the modern workplace suggests that some of the choices in workplace design, intentional or not, sometimes come at a cost to individual productivity. The
psychological literature on attention, cognitive load and task switching explains the limits of human attention and explains how these environments challenge our abilities. Work in the economics literature on time discounting and impatience explains how people may be tempted by short term immediate gratification and may neglect less enjoyable activities that have long term benefits. HCI studies on disruptive software applications provide insight into how our limited patience and attention are challenged by particular push technologies.

Section three describes two studies about how people manage interruptions during work. The first investigation, a four-group between-subject experiment investigates what switching behaviors are like, how much people allow themselves to be distracted by alternative tasks and the impact the distractions have on their task performance. Participants answered two sets of questions, the first of which was tedious but paid well and the second of which was more enjoyable but paid little. Depending on condition, the participants either completed all of one category or the other, or else have the option to switch back and forth between types of problems.

The second study examines one particular potential disruption from work, namely, email. This survey describes power users experience with email, their understanding of the problems associated with it, and whether email is disruptive to work. It presents how two groups of email users who receive
large quantities of mail switch to and from email throughout the day and demonstrates how fractured their workday becomes. It also demonstrates that people may not have a strong sense of how to be efficient with email tasks.

Section four of this thesis presents how people using modern computers actually switch between tasks and discusses what can be determined about their productivity from watching their switching behavior. A survey of two groups of people's work habits is used to describe how people report their work habits in general. This investigation examines what people do with their computers when they are working hard, working less hard, and not working. Another study on how people work with their computers in real world settings is then described. In this study, the Worksniffer software created for this task records when people switch from one application to another and how much they type in each application. This software was installed on participants' computers and participants were asked to describe what they were doing while using the computer. The results suggest that it is possible to identify both general and individual patterns that describe when people are working and not working.

The prototype and implications are then presented. Simply Work, is a visualization of a user's task switching and work process. The application is built upon the WorkSniffer application used in the real-world study. This tool assists users in self-monitoring their work habits by displaying information
about how much they switch between applications and how much they type. The hope is that this tool will help users to maintain deep focus, reflect upon their work practice, and take breaks in a timely manner. Finally some principles on how this information can be used to design both push applications and applications that encourage self-reflection upon individuals own work habits are presented. These ideas include general thoughts on what user interface designers might consider, what software can tell about peoples work habits from their switching behavior, what organizations might tell their employees, and what other tools might be built for self-reflection.

Ultimately the goal of this thesis is to shed some light upon how people work with their computers, what is successful and what is not, in order to provide some contribution to the understanding of how computers can better be designed for the capabilities of those that use them.

2 Background Literature

2.1 Attention, cognitive load and task switching

Psychologists work on cognitive load and dual tasks has revealed much about the limits of our attention. The active memory that we use to manipulate information is called short-term or working memory (Miller 1956.) This type of memory is broken into chunks, units of information (Hayes 1989). We can only keep 7 plus or minus two chunks active in our working memory at one time (Miller 1956), and always work within these basic limitation of memory.
Cognitive load, from this standpoint, is the amount of activity in working memory at a given time. If we are working on multiple tasks we will fill up our working memory quickly, and have to transfer different chunks in and out as we switch tasks. When we are in these high-load situations (for instance when we are switching between tasks or are interrupted), our performance on both tasks may degrade.

All interruptions are not equally disruptive. Interruptions that are similar to the main task and require immediate attention (as email sometimes may be) seem to be more disruptive (Gillie and Broadbent 1989.) If people can negotiate when an interruption occurs, they perform better on their primary task, but they may not complete the interrupting task in a timely fashion (McFarlane 1999.) This is reassuring since many interruptions during computer work are negotiable. And the cost of switching increases as the complexity of the rules needed for performing the tasks increases (Rubinstein et al 2001) as is the case with many computer tasks. Thus, the limits of our attention may be ill suited to the ease of switching and the amount of incoming information characteristic of modern computer use.

2.2 Disruptive software applications

HCI researchers have investigated the cost of interruptions that are specific to personal computers, such as email, instant messenger and web surfing. Miyata and Norman (1986) first reviewed the work on attention and dual-tasks and
specified three phases of computer tasks—planning, execution, and evaluation—during which being interrupted would have different effects. Since then several researchers have looked at how the features of the interruption and the primary task impact the time it takes to complete both the interruption and the task. Consistent with the findings on the traditional attention literature, Cutrell et al (2000) interrupted instant messenger users during these different phases and found that being interrupted during evaluation resulted in significantly longer completion times than during the other phases and that interruptions with irrelevant information took longer to recover from. Bailey et al (2000) also found that the type of work (such as addition, reading comprehension, image comprehension) impacted the amount of time it took to complete a task after an interruption.

Other researchers have observed real-world switching behavior. Jackson et al (2001) focused on email usage and found that it takes on average one minute and 44 seconds for employees to react to a new email confirmation! In addition, it took an average of 64 seconds for them to recover and return to their work after they finished with the email. Given this number, a person who receives 50 emails a day one at a time and switches to read each one would experience almost an hour in recovery time alone from email. Thus it is just as important to understand how people choose to switch tasks and not just the limits of their attention.
2.3 Time discounting and impatience

Even if we have some understanding of cognitive load and the cost of switching, we may not adjust our work habits accordingly (as the author can certainly attest to!) The short-term benefits of pleasurable activities are clear: we enjoy them. Activities, such as work, that are not fun but are important, are harder for us to envision while we can see the costs of work today clearly. What is in front of us is immediately salient. What will happen later is not.

Philosophers, psychologists and economists alike have documented this mode of operation throughout the years. (cf. Ainslie and Haslam). Behavioral economists call this time discounting. The further from now a reward will be received, the less its perceived value is. This idea, which has been used to explain many behavior sets from addition to procrastination, is useful to unpack why we may switch tasks even if we know we shouldn't. It may be hard to see the long-term benefit to staying on task, especially if the switches feel short and transient. But from the attention literature, we know that, in fact, these switches do have negative impact upon our work.

3 Two Studies On Task Switching

It is important to understand not only what people can reasonably attend to when faced with dual-tasks but also how they choose to switch. This section discusses two studies that investigate this issue. The first is an experiment on task-switching in which people can choose whether they should work on a
quick, fun variable but low-paying task or a more tedious, high paying one. The second study looks at how two user groups who rely heavily on email manage and perceive it.

3.1 Experiment on Task Switching

This experiment addresses whether people switch between tasks efficiently and how costly their switching is to their productivity. A four-group, between-subject experiment was run in which the experimental groups were able to switch between two problem solving tasks at will whereas the control groups could not. I hypothesize that time on task, performance, and switching behavior are negatively impacted by the ability to switch between tasks.

3.1.1 Methods

Thirty-two participants were recruited through on-campus flyers at MIT. Participants included undergraduates, graduate students, researchers, administrators, alumni and business people. After reading and signing the consent form and completing a brief demographic questionnaire, participants were told they were going to solve puzzles on a computer and the procedure was explained (Appendix 1).
This application was built using Macromedia Flash and Visual Basic scripting. All participants solved the same set of 5 primary task problems and ten secondary task problems (see Figure 1.) For the primary task problems subjects were asked to add up multiple numbers until they reached the sum of 100, or if that was too challenging, any number between 80 and 120. For each problem there were eight to twelve numbers in a set and finding the total of 100 required between three and five numbers. The numbers were added up for the participants as they selected boxes next to each number. Still these problems were difficult because the problem space was very large.

The secondary task problems were alphanumerical sequences or sets of items in which the participant had to select which additional item belongs in the set. Payment was depending upon how they solved these questions. Responses were either multiple choice or fill in the blank. Participants in post-experiment
interviews described some of the secondary problems being harder than the primary problems and some easier.

![Figure 2: Unseen Control Condition with Add to 100 Problem Displayed](image)

Payment was dependent upon task performance. For each primary task problem solved correctly a participant earned either one dollar for adding the numbers up to 100 or 25 cents for adding the numbers up to any number between 80 and 120 except for 100 (see Figure 2.) The secondary problems were worth on average 25 cents. Participants were told some would be worth more than 25 cents and some less but they were not told exactly how much the problems were worth before solving them. All the secondary problems were worth five cents except one that was worth $2.05. When they selected the problem, they saw the value of it, along with the problem itself.
In the two control conditions participants either a.) submitted an answer for each of the primary task problems first and then the secondary task (the add to 100 first condition) or b.) submitted an answer for the entire secondary task first and then the primary (the sequence completion first condition). In the first experimental condition, the secondary task problems appeared while the participants were working on the primary problems (the seen experimental condition.) In other words, when the participant first started the application, no secondary problems were visible. As they worked, one appeared and then another, etc. In the second experimental condition, the secondary task problems appeared while the participants worked on the primary task problems, but they had to click a button make the problems visible (the unseen experimental condition.) Participant could switch between problems as they wished, but once they submitted their answer, the associated problem would disappear. Subjects knew that they were limited to 20 minutes to complete as many problems as they could.

A log file recorded every time a participant looked at a problem, and if they submitted an answer, what they answered. Every action was recorded with a timestamp to the second.

3.1.2 Results

Three sets of variables were analyzed: amount of time on task, number of correct answers, and number of switches. All analyses were unpaired t-test.
3.1.2.1 Time to Complete Tasks

A significant difference in total time on all problems was found between the experimental *seen* condition ($M=701$, $SD=170$) and the *add to 100 first* control ($M=934$, $SD=152$), $t(14)=-2.89$, $p <.01$. and also the *sequence completion first* control ($M=926$, $SD=167$), $t(14)=-2.67$, $p<.01$. The data is tending towards significance on the total time for all problems between the experimental *unseen* condition ($M=719$, $SD=344$) and both the *add to 100 first* ($M=934$, $SD=152$), $t(14)=-1.614$, $p =.13$ and *sequence completion first* ($M=926$, $SD=167$), $t(14)=-1.528$, $p=.15$ controls. When one outlier that was more than two standard deviations away was removed from the *unseen* experimental condition, a significant difference on total time on all problems was found between the experimental *unseen* condition ($M=609$, $SD=164$) and both the *add to 100 first* ($M=934$, $SD=152$), $t(13)=-3.981$, $p<.001$ and *sequence completion first* controls ($M=926$, $SD=167$), $t(13)=-3.699$, $p<.002$.

A significant difference in time on the *add to 100* problems was found between the experimental *seen* condition ($M=399$, $SD=162$) and the *sequence completion first* control ($M=573$, $SD=156$), $t(14)=-2.184$, $p< .05$, but not the *add to 100 first* control. A significant difference in time on the *add to 100* problems was found between the experimental *unseen* ($M=313$, $SD=158$) condition and both the *add to 100 first* ($M=572$, $SD=156$), $t(14)=-3.314$, $p<.005$. and *sequence completion first* ($M=519$, $SD=225$), $t(14)=-2115$, $p<.05$ controls.
There were no significant differences on the time spent on the sequence completion problems.

Note that the direction of all effects was the inverse of what was predicted: those in the experimental conditions took significantly less time than those in the control conditions.

3.1.2.2 Correct Answers
There were no significant differences on total number of correct answers, number of correct sequence completion answers, or number of correct add100 answers, whether including only perfect responses (added up to 100) or both perfect and imperfect responses (added up to any number between 80 and 120).

3.1.2.3 Amount of Switching
When all switches were counted (including switching within one set of questions for both the control and experimental conditions and across types of questions for the experimental conditions) and also when only switches within types of questions were counted, there were no difference in the total number of switches between any conditions.
3.1.3 Discussion

In this experiment, the ability to switch tasks did impact how quickly people solved problems, but in the opposite of what was predicted. People who could switch between tasks actually solved the problems more quickly without a cost to their overall performance. Why did this occur?

One explanation is that people managed their time better when they could switch. They took timely breaks. Switching then may have reduced interference between successive iterations of the same task, as well as possibly giving subjects a chance to recover from mental fatigue.

Also there may not have been dual-task interference as expected. Both tasks were easy to understand. Participants may have been able to reestablish context without increased cognitive load. One would only expect performance differences if the experimental groups were highly cognitively loaded and seeing other problem pop-up or knowing they would pop up or switching out of frustration and losing time in the switch caused their performance to degrade.

A final post-hoc explanation involves the impact of affective state on performance. Informal discussion with control condition participants revealed that not knowing how difficult the next task was or how much time it would take was stressful and unpleasant. That finding, coupled with the experimental outcome, suggests a similarity between this experiment and many done by
Alice Isen. She and her associates have used methodologies in which participants in different experimental conditions perform the same tasks as each other, but with slightly different affective states. They have found that time on task degrades when negative affect is induced and improves when positive affect is induced (cf. Estrada, Isen & Young 1998). It is quite possible that control subjects were more frustrated and therefore performed more slowly than those who could switch.

It is worthy of noting that one reason mathematical and logic puzzles were used as a manipulation was because the student population, who comprised a large percentage of the participants, enjoyed them and were very good at solving them. Earlier versions of the study involved easier questions, which were too easily solved. Therefore in this version the problems were more difficult. It seems that even with the increased difficulty of the problems there may have been a ceiling effect on task performance.

3.1.4 Conclusion

In the future better competition between the two tasks might reveal findings that correspond with the predictions in the current study. This may be due to limited cognitive load between tasks or to an unintended negative affect manipulation in the control conditions.
3.2 How Efficiently Power Users Manage their Email

For many computer workers, email is a critical tool for managing relationships, communicating professional needs and organizing work throughout the workday. Indeed some individuals manage so much email that recently Deborah Fallows of the Pew Internet and American Life labeled this group, power users (2002). Most of her respondents reported that they receive little email and are not disturbed by the amount they receive. However, the power users apply email to the widest range of tasks (such as arranging appointments, managing documents, asking questions about work) and reported being more accessible than non-power users, that the email load posed more of a problem than non-power users, and that they approach their email as necessary chore. These findings suggest that though power users appreciate email, they recognize that it comes at a cost. And based on the Jackson et al.’s findings how long his participants to recover from a switch (2001), this cost may be greater than the respondents believe.

Yet users may be distracted by email may if they switch in a way that is either not efficient or based a real need to read and respond. I propose that when people switch from other applications to email, they underestimate the time it will take to return to their previous task and also that people have a hard time resisting the temptation of reading email.
The current investigation examines power users' experience with email, their understanding of the problems associated with it, and whether email is disruptive to work. It first discusses how much email the participants receive, how frequently users read and respond, and how much time they spend on it as an attempt to quantify how much of the participants' workdays are spent on email and also how fragmented email checking might make it. Then it investigates what percentage of the email people receive is perceived as urgent. If a lot of email is urgent, it makes sense to check often and, perhaps, spend a lot of time on email. Finally it investigates what the participants feel is an efficient way to check their email and whether they do it.

3.2.1 Methods

Two separate surveys were conducted. For the first survey, 228 MIT undergraduates from computer science and electrical engineering were recruited through a departmental mailing list. Participating students were entered into a lottery to win a gift certificate to the campus store. For the second survey, 43 Media Lab corporate sponsors and 5 lab members were recruited from a laboratory mailing list. These participants were unpaid.

In both cases participants filled out an online survey that asked them to identify how many emails they received a day, how much time they spent on email, whether their email client checked for email, how often the client or they checked for email, how often they read their email, how often they
responded, the percentage of email that was urgent, what time schedule of checking email was the most efficient and if they didn’t check their email in the most efficient way, why not (see Appendix A). In addition, laboratory participants were asked what their affiliation with the lab was (laboratory sponsor versus employee.)

3.2.2 Results

3.2.2.1 Amount of email received

Knowing the amount of email someone receives is one basic measure of how disruptive it may be for them. If they receive very little email, it is likely that they are undisturbed by it. Of course, if they receive many emails, this does not necessarily mean it is a problem. But it would suggest a need for further inquiry.

Respondents reported how much email they received using a 6 point scale: less than 10, 10-25, 25-50, 50-100 100-500 and greater than 500. For each category of response on an average value was calculated, e.g., the category 10-25 was scored as 17.5. The score for >500 was counted as 600. (This same method of finding the middle value and using it to find an overall mean is used to compute means throughout this email section.) Computed in this way the mean number of emails received for sponsors was 51 and for students was 75. Eighty percent of laboratory sponsors reported that they receive between 25 and 100 emails daily. Seventy-three percent of student reported that they receive
between 10-50. Both groups report that they receive a substantial amount of
e-mail daily (see Figure 3.)

![Bar chart showing number of emails received daily]

**Figure 3**

3.2.2.2 How often they check e-mail

These two populations receive a large number of e-mails a day. But how often
do they actually stop what they are doing to check their e-mail? If they do it
often, it is more likely that their e-mail is taking them away from other tasks.

Participants were asked whether they had an e-mail client that checked their
e-mail and how often it was set to check for e-mail, and if their client didn’t
check, how frequently they themselves checked. Based on informal feedback
from participants, it appears that the phrasing of these questions was
confusing. Also it is possible that some participants were unaware whether and
how often their e-mail clients checked for mail. Both of these points should be
taken into consideration when reviewing the findings in this section.
Respondents reported how often they checked their email using a 7 point scale: 0-5 minutes, 5-15 minutes, 15-30 minutes, 30 minutes to 1 hour, 1 to 2 hours, 2 to three hours, and over 3 hours. Twenty-seven percent of sponsors and 39 percent of students reported that their email client checked for mail automatically (see Figure 4.) For sponsors forty percent of their clients checked between once and twice an hour and for students about 20 percent reported for each of the time intervals over the 15 to 30 minute time interval. For the participants who checked themselves, 53 percent of sponsors and 44 percent of students checked between every 5 and 15 minutes.

![Figure 4](image)

3.2.2.3 How often they read email

In the current survey participants were asked separately how often they checked or their software checked for email and also how often they read the body of the email, which is what is reported in this section. Checking email
may distract us only for a moment, but reading email distracts us longer and there is likely a longer switch time cost as well. Respondents reported how often they read their email on a 7 point scale: 0-5 minutes, 5-15 minutes, 15-30 minutes, 30 minutes to 1 hour, 1 to 2 hours, 2 to three hours, and over 3 hours. Sponsors read email every 98 minutes and students every 52. 53% of sponsors read their email every 5 minutes and 83 percent read it within every 30 minutes (see Figure 4.) Similarly, 44 percent of students read email within every 5 minutes and 74 percent read within every 30 minutes. Using the same method of converting the categorical data to average values as before, the mean value of time spent on each email for students was two minutes and for sponsors 4.3 minutes.

![Figure 4](image)

But reading email, which here seems to happen quite often, is only part of email management. Sometimes you have to respond.
3.2.2.4 How often they respond to email

One could argue that reading email is fun but and writing is tedious and time-consuming. Responding to email likely takes up more time than reading. How often do these two groups respond to email?

Respondents reported how often they responded to their email using a 7 point scale: 0-5 minutes, 5-15 minutes, 15-30 minutes, 30 minutes to 1 hour, 1 to 2 hours, 2 to three hours, and over 3 hours. The sponsors responded to email every 56 minutes and students every 97. Even though they read and receive email a lot, 53 percent of students respond rarely: less than every 3 hours. Sponsors describe their response rate to email quite differently in that there is a greater spread to how often sponsors respond. Twenty-eight percent respond between every 5 minutes to 15 min, 18 percent respond every 30 minutes to one hour, and 22 percent respond every one to two hours (see Figure 5.)

Figure 5
3.2.2.5 Amount of time on email

It appears that both populations check their email at astounding rates and sponsors respond often while students do not. But if this checking takes little time, then we need not be concerned that email cuts into their workdays. This next section asks how much time is spent on email.

Respondents reported how much time they spent on email using a 7 point scale: 0-5 minutes, 5-15 minutes, 15-30 minutes, 30 minutes to 1 hour, 1 to 2 hours, 2 to three hours, and over 3 hours. Sponsors spent on email 71 minutes on average while students spent 53. About 45% of sponsors spend 1 to 2 hours on email every day and 28 percent spent greater than 2 hours (see Figure 6.) Students spent a wide range of time but 85% spent 15 minutes to 1 hour. Still some spent very little (10 percent spend less than 15 minutes) and few spend a lot (only 6 percent spend over 2 hours.) Students may spend less time on email than sponsors since they respond so infrequently.

Sponsors spent a mean of 4.2 minutes on each email they receive and a median of 1.6 with a standard deviation of 8.22. For students the mean amount of time per email was 2.0, the median was 1.3 and the standard deviation was 3.4. This seems like a reasonable amount of time to spend on email. However, it is possible that the amount of time on email was under-reported. If people are constantly reading their email, as the previous section suggests, they may be using email as a means of taking a break during intense work. (Informal
personal communication with these populations and others support this speculation.) If email is used as a distraction when work is difficult, users may not be accomplishing as much as they might without email and also they may find it difficult to accurately report this small time intervals.

![Figure 6](image)

**Figure 6**

3.2.2.6 Percent of email that is urgent

If people believe that a great deal of their email is urgent, they should read it and respond often. But if they don't believe much of their email is important, they should read and respond infrequently. Since these two populations do spend a good amount of their days on email, one would hope that a great deal of their email is urgent.
Respondents reported how much of their email was urgent using a 6 point scale: 5%, 10%, 25%, 50%, 75%, and 100%. On average sponsors rate 16 percent of their email as being urgent while students rate 12 percent urgent (see Figure 7.) 70 percent of sponsors and 80 percent of students say that around 10 percent of their email is urgent. Thus they believe very little of their email requires immediate attention. But if so little email is urgent, then it may not make sense to spend so much time on it. It is possible that these respondents do not manage their email efficiently as they might. Still, if these small numbers of email are extremely important, it may be worthwhile for the respondents to check their email regularly. The next section addresses whether these two groups of email users may answer this open question by addressing what they think is the most efficient way to check email.
3.2.2.7 The most efficient way to read email

These participants spend a lot of time on email though little of this mail is urgent, which seems inefficient. But what do they think? Do they believe that they read email in an efficient manner?

Respondents reported what they thought was the most efficient way to read their email using a 7 point scale: 0-5 minutes, 5-15 minutes, 15-30 minutes, 30 minutes to 1 hour, 1 to 2 hours, 2 to three hours, and over 3 hours. Sponsors reported it was most efficient to read email every 53 minutes on average and students reported 135 minutes on average (see Figure 8). Seventy-three percent of students believe that email should be read once an hour or less and 32 percent believe that email should be read every three hours or less. Students firmly believe that email should be read infrequently than they actually do.

![Figure 8](image)
For sponsors the spread in responses was greater. Forty-seven percent believe that the most efficient way to read email between every half hour and two hours whereas 36 percent believe it is best to read between every five and thirty minutes.

However, in both cases, some of the respondents who read their email most frequently express a belief that to read far less frequently is preferable and a few seem to think reading much more frequently is most efficient.

3.2.2.8 The Relationship between Perceived Efficiency and Reading and Receiving of Email

With some exceptions, these two populations overall feel that they check their email in an fairly efficient manner though some believe that checking less often is preferable and a few think more often is better. What do these populations base their assessment upon?

The data was coded as reading more often than what is perceived as most efficient, reading less often, and reading the same amount. Forty-two percent of sponsors and thirty-nine percent of students believe that they read their email the most efficient way whereas thirty-three percent of sponsors and eighteen percent of students feel they read too often and twenty-five percent
of sponsors and forty-three percent of students feel they read not often enough. This same pattern is seen in the following plot of the difference between the reported reading rates and the perception of most efficient reading rates of both students and sponsors. Positive numbers indicate the amount they feel they read too much and negative numbers indicate the amount they feel they read too little.

Many sponsors and students seem to believe they are reading the right amount. It is possible that these groups read their email in the most efficient way, however, it seems more likely that they base their evaluations of efficiency upon their own behavior. There seems to be a trend that some sponsors feel they should read less often and students feel that they should read more often.
3.2.3 Discussion

Previous work suggested that only the most engaged email users believed that email had serious drawbacks and even they saw great benefit to the medium as well. Still, email users tended to read their email often and switching back from email came at a cost to the first task at hand.

In this study, two groups of computer users who receive a large amount of email daily describe their experiences with the communication medium. Both groups, but the sponsor group, in particular, read their email with amazing frequency, suggesting that email likely distracts them from other tasks at work. If so few email messages are urgent, as both groups reported, perhaps it is not in the users interest to read as much as they do and certainly it is not in the sponsors best interests to respond as much as they do. Students may read and respond less often than sponsors simply because email means something very different to them. Email may be work for students rarely and when it is, a response within a day may be fast enough. On the other hand, sponsors may respond quickly because communication is a key part of their work and a fast turn-around is expected. In fact, their image at work may, in part, be based on their responsiveness and how quickly they respond to email may be one facet of this measure. But if this expectation is based on a sense that everything is urgent, even problems that are not important, and it interferes with the sponsors work, the speed of response required in business may actually do a disservice to the productivity (and perhaps, well-being) of its workers.
One could reasonably consider it encouraging that an overall trend was found between how the participants read their email and what they felt was the most efficient way to do so. They may be very efficient at checking email. On the other hand, they may simply be saying that their own way is most efficient without giving it more thought.

Two final caveats: this investigation relies upon participants self-reporting of their experiences. It is possible that particular features of their email usage may be under or over exaggerated. People may exaggerate to make themselves appear more favorably. They may not remember junk mail or may not count certain messages such as news and mailing lists. Observation of real-life email behavior would be an excellent complement to the current study. Second, this paper assumes that email takes away from work time. But email may instead replace other means of communicating and taking breaks. Email may even save time if it helps people to maintain relationships and take breaks more efficiently. Still, this does seem like an unlikely explanation given the amount of task-switching these participants described.

3.2.4 Conclusions

Though previous research is split upon whether email is often a distraction for intense users, this study found that some users spend a good part of their day using it and distract themselves from their current work to manage mail. This
rapid switching likely comes at a great cost to the individual’s productivity, if not their satisfaction with their workday.

Since these groups of users switch often do not have a strong sense of how to determine how to make their email usage more efficient, designers of email clients might want to take the cost of switching to and from email into account in their interface designs. The findings of this study also suggest that more research into describing efficient email checking is warranted. Some specific ideas about design of push applications and future research are discussed in the last section of this thesis.

4 System Development

4.1 Investigations into Work Habits

By closely observing how people work with their computers, we may be able to 1.) learn when they are working efficiently and when they are not and 2.) understand when interfaces support efficient work and when they don’t and 3.) how interfaces can be improved. This next section describes work on software that tracks how people work with computers and makes assessments or recommendations to better support their activities, particularly through self-reflection upon work habits.
4.1.1 Some Existing Work

Researchers have used a variety of techniques to track users work and also to provide feedback about it. Lockerd and Mueller’s Cheese project tracked how web users moved their mouse and then created a model of the user (2001). Microsoft Research is working on several systems that not only track but also use the information to try to adapt to the user’s current and long-term needs. I’ll mention here only two from their Adaptive Systems and Interaction Group’s suite. Lumiere, a compelling user modeling system project that lead to the dreaded Microsoft paperclip, used a Bayesian network to decide what a user is doing by looking at how the person was using different software applications currently and through a persistent model of the user’s expertise (Horvitz et al 1998). LookOut mixed-initiative system parses text about scheduling appointments from email messages and automatically opens Outlook’s calendar and fills in what it knows about the appointment (Horvitz et al 1990). The user can then accept, modify, or reject the appointment.

This previous work does not focus on how people switch tasks with their computers, which may well be an important problem. The following two investigations consider how people switch tasks, whether they seem to switch well, and whether they appear to have insight into the cost of switching. The first is a survey on how people use different applications when they are working hard versus when they are using their computers, but not for work.
4.1.2 Survey on People’s Work Habits

To understand when people switch tasks while working with their computers, one must first identify what is the unit of a task and when a switch occurs. One very basic way to identify switching is to look at switches between computer applications. However, not only must we know that people are switching, we also need to know whether the task being switched to and from are work-related. Is it possible to label certain applications as being used primarily for work or not for work? In this investigation, two groups of computer users were asked about their habits using a computer while working and not working and some interesting patterns were found.

4.1.2.1 Methods

One hundred and seventy one MIT undergraduates from the computer science and electrical engineering department and 83 laboratory sponsors were recruited through mailing lists. Participating students were entered into a lottery to win a gift certificate to the campus store whereas sponsors were unpaid.

Participants filled out an online survey that asked them to identify what application they had open both when they were working hard without distraction and also when they were using a computer not for work (Appendix 7.4). They were then asked to rank in order how much they used the applications that were open under both conditions.
4.1.2.2 Results
Some applications were described by participants as being used primarily for either work or not for work (Table 1). For instance, spreadsheet, programming statistics, and mathematical programming applications were almost never used for anything but work for these user groups. Word processing was used for work often and only a small amount for not work.

<table>
<thead>
<tr>
<th>Percent of Respondents Who Had Application Open in Different Work Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Word Processing</td>
</tr>
<tr>
<td>Email</td>
</tr>
<tr>
<td>Spreadsheet</td>
</tr>
<tr>
<td>Programming</td>
</tr>
<tr>
<td>Web</td>
</tr>
<tr>
<td>Statistics &amp; Math</td>
</tr>
<tr>
<td>Games</td>
</tr>
<tr>
<td>Multimedia Editing</td>
</tr>
<tr>
<td>Music Listening</td>
</tr>
<tr>
<td>Web Editor</td>
</tr>
<tr>
<td>Organizer</td>
</tr>
</tbody>
</table>

Table 1

Games and multimedia authoring was used by both groups primarily not for working. Though these users report that they listen to music much more often while not working, many listen while working as well. Email, web and
organizers were all used in both categories and cannot be used as a means of distinguishing work from non-work for these users.

It appears even more can be determined about when people are working by looking at which applications were open only for working, only for not work, or for both or neither conditions (Table 2).

Some applications were open predominantly for work only including programming, spreadsheets, statistics and mathematic applications. Games was the only type of application opened predominantly only when not working. Other applications were used either for work or for both work and not work such as word processing, web and email. Still there were differences between which applications the students had open versus the sponsors.
Percent of Respondents who Had Applications in Different Combination of Work Conditions

<table>
<thead>
<tr>
<th></th>
<th>Open for work but not outside of work</th>
<th>Open outside of work and not for work</th>
<th>Open for both</th>
<th>Open for neither</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Student</td>
<td>Sponsor</td>
<td>Student</td>
<td>Sponsor</td>
</tr>
<tr>
<td>Word Processing</td>
<td>64</td>
<td>66</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Email</td>
<td>6</td>
<td>16</td>
<td>23</td>
<td>2</td>
</tr>
<tr>
<td>Spreadsheet</td>
<td>96</td>
<td>41</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Programming</td>
<td>55</td>
<td>21</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Web</td>
<td>5</td>
<td>16</td>
<td>21</td>
<td>14</td>
</tr>
<tr>
<td>Stats and Math</td>
<td>15</td>
<td>5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Games</td>
<td>0</td>
<td>1</td>
<td>45</td>
<td>16</td>
</tr>
<tr>
<td>Multimedia Editing</td>
<td>4</td>
<td>19</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>Music Listening</td>
<td>3</td>
<td>1</td>
<td>45</td>
<td>29</td>
</tr>
<tr>
<td>Web Editor</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Organizer</td>
<td>5</td>
<td>30</td>
<td>10</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2
It is useful to know what application people have open when they are working but it is even more important to know what applications they are actually using. Table 3 presents how many participants ranked particular applications in

<table>
<thead>
<tr>
<th></th>
<th>Number of Respondent Who Ranked Applications Differently for Working and Not Working and the Average of the Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of Responses (N=171)</td>
</tr>
<tr>
<td><strong>Student</strong></td>
<td><strong>Average of Responses</strong></td>
</tr>
<tr>
<td><strong>Number of Responses</strong></td>
<td>39</td>
</tr>
<tr>
<td><strong>Email</strong></td>
<td>147</td>
</tr>
<tr>
<td><strong>Spreadsheet</strong></td>
<td>37</td>
</tr>
<tr>
<td><strong>Programming</strong></td>
<td>37</td>
</tr>
<tr>
<td><strong>Web</strong></td>
<td>138</td>
</tr>
<tr>
<td><strong>Stats and Math</strong></td>
<td>32</td>
</tr>
<tr>
<td><strong>Games</strong></td>
<td>41</td>
</tr>
<tr>
<td><strong>Multimedia Editing</strong></td>
<td>34</td>
</tr>
<tr>
<td><strong>Music Listening</strong></td>
<td>102</td>
</tr>
<tr>
<td><strong>Web Editor</strong></td>
<td>36</td>
</tr>
<tr>
<td><strong>Organizer</strong></td>
<td>35</td>
</tr>
</tbody>
</table>

**Table 3**
both the working and not working conditions and the average difference between the ranks in the two conditions. (Participants only ranked applications that they used in a given condition so in many cases applications were not ranking both for work and not for work.) A negative score indicates that the particular application was used more for work whereas a positive one indicates it is used more not for work.

Some applications were used more both for working and not for working than others. Web browsers and email were used by both students and sponsors for working and not for working. Many students also listened to music under both conditions. A greater percentage of students used more applications than sponsors. Students played games, listened to music, edited multimedia and to a lesser extend used email much more when not working. Sponsors also played games, and listened to music but also used their web editors and to a lesser extent the web when not working. It is interesting that students actually rank web usage similar for work and not for work whereas sponsors do the same with email. Students may use the web for research on homework assignments as well as for having fun whereas sponsors may collaborate and organize via email as well as sending personal messages.

Across the two user groups, some applications such as statistics, math and programming were used predominantly for work while others were used predominantly outside of work such as games. This suggests that some idea of
when people are working and not working can be distilled from observing which applications they have open and how much they use those applications. For some critical applications such as web browsers and email clients, it is likely the story cannot be uncovered without knowing something about the content that people are writing or reading.

The two user groups might have very different computer use patterns when working versus not working. It is encouraging that these users’ descriptions of their own work habits might help systems that identify which applications are used for work and which are not. One could imagine user groups that would definitely have different patterns for applications being open for work versus not for work. For instance, designers might have multimedia authoring tools open while working. And a businessperson who yearns to be a writer might use a word processing for personal pursuits. Individual users likely have work patterns that systems could likely identify with some ease. The more difficult part would be interpreting when someone is working or not, when their given state of working is consistent with what they wish to be doing, and how to best support the user’s desired work habits.

It is interesting to know how people report upon their work, but their reports might be quite different from what they actually do. The next study analyzes how people work on real-world tasks.
4.1.3 *Observing Task Switching in Computer Work*

The previous survey on work habits asked people to report upon what applications they use and how much when they are working and not working. Still, this information needs to be related to how people work in real life in order for its findings to demonstrate that it can be used to help identify real work habits. By looking at what applications people have open, how much they use them and how much they switch between applications while they work, the following study provide insight into how people manage multiple tasks in the real world. The goal of this section is to use minimally invasive monitoring software to identify particular patterns of when and how people are working. Designers can use this information to build interfaces that support the best of habits and to discourage the worst. For instance, interfaces could know how and when to interrupt or could integrate applications that are often used together.

This study uses an application, *WorkSniffer*, that logs people’s work habits. Built using C++ and Windows API calls, this application, which is invisible to the user, logs when people logon to a machine, when they open applications, when they switch between applications, and how much they type alphabet characters, numeric characters, shift, control, up, down, left, and right (Appendix 7.5).
4.1.3.1 Study

In this investigation 4 graduate students and 2 post-doctoral appointees installed the WorkSniffer application on their laptop for one workday. At the end of the day they completed the same work habits survey described previously, then sat down with a printed copy of the log file and this researcher to annotate what they did during the day. The annotation began with the experimenter asking the participant to rank on a 7-point Likert scale how productive their day was compared to other days. The participant was told that this ranking should be compared with other days as a graduate student and that productive should be thought of as getting their research done and working towards graduating. As many of the participants were either taking general exams or writing their theses, their short-term current research goals were well defined. They were then to describe what they did that day: both work accomplishments and distractions or non-work activities. Next the participant reviewed segments of their day with the experiment. A segment was generally greater than a half hour of time and the bounds of the segment was determined by when the participant changed from one set of goals to another. So, for instance, a participant might start her day by catching up on email—one segment—and after completing that move on to gathering articles for her general exams—a second segment. For each segment participants were asked to describe what they did in that time segment by considering what applications were open, how much they were used, and how much they switched. For each segment, the participant ranked their productivity the same
way as before. As the participants described their segments, the experimenter identified which applications were used for which tasks if it wasn’t obvious (such as web browsers that were used to check personal mail or unusual application that the WorkSniffer software did not easily identify.) Finally the participants were asked if they wanted to change their assessment of how productive they were after going through their day with the experimenter. The experimenter emphasized that she did not care whether they felt they should change or not change their assessment, but she was only giving them an opportunity to reassess their behavior after the discussion about it.

The data was coded in two basic ways. First, by switch. Whenever a user moved from one application to another, what they were doing with that application—such as how much they were typing and how long they used the application—was coded. (A switch here is different than a switch in the cognitive load literature and from how I have been previously discussing switching. These switches are between applications and many applications may be used for one task. For instance, if someone is building a reference list, they may be copying sources from their email inbox or from the web to EndNote or Microsoft Word.) Second, the data was analyzed by segments, which, as described previously, correspond to a unit of time in which a particular or set of tasks was accomplished. Both of these datasets were then coded two ways: on aggregate and individually.
4.1.3.2 Results

In the previous work habits survey, spreadsheets, word processing, programming, and statistics/mathematics applications were all associated with work. The current set of participants responded similarly (see Table 4).

<table>
<thead>
<tr>
<th>Participant Description</th>
<th>Used Only for Work</th>
<th>Used For Both</th>
<th>Used Only Not For Work</th>
</tr>
</thead>
<tbody>
<tr>
<td>HCI researcher who does not program.</td>
<td>word, spreadsheet, graphics</td>
<td>email, web, music</td>
<td></td>
</tr>
<tr>
<td>HCI researcher.</td>
<td>word, programming</td>
<td>email, web, music</td>
<td>graphics</td>
</tr>
<tr>
<td>Programmer.</td>
<td>word, windows explorer</td>
<td>email, web, music</td>
<td></td>
</tr>
<tr>
<td>HCI Researcher.</td>
<td>word, programming</td>
<td>web, music</td>
<td>email</td>
</tr>
<tr>
<td>HCI researcher who programs.</td>
<td>programming, spreadsheet, word.</td>
<td>email, music, web</td>
<td></td>
</tr>
<tr>
<td>Psychologist who does not program.</td>
<td>word, statistics, spreadsheet</td>
<td>email</td>
<td>web, music</td>
</tr>
</tbody>
</table>

Table 4

Note that participants were not asked to respond on how they used instant message clients and the programmer wrote in Windows Explorer. All the HCI researchers only program for work whereas the programmer also does it outside of work. Most participants use email and web for both work and non-work. The psychologist categorizes both web and music as non-work applications, which no one else does. With such a small sample it is impossible
to tell whether these variations are individual differences or preliminary evidence that different professions have different styles of application use.

From this data it appears that though there were many shared characteristics of working versus not working, participants may have some individual differences in how they use particular applications. One way to further investigate this possibility is to compare participants segments that they rated most productive to the ones they rated least productive. Two sets of switching data representing participants most and least productive segments was selected for each participant. A series of t-test comparing individuals most and least productive time segments revealed interesting results.

A significant difference on number of switches was found between individual ratings of their least and most productive section. This difference was significant in both directions depending on the individual (see Table 5). In other words, some participants switched significantly more when they were working the most and some switched significantly less.

<table>
<thead>
<tr>
<th></th>
<th>Mean Diff</th>
<th>DF</th>
<th>t-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>L:H Total</td>
<td>.002</td>
<td>643</td>
<td>.551</td>
<td>.9696</td>
</tr>
<tr>
<td>L:H 1</td>
<td>.097</td>
<td>85</td>
<td>.373</td>
<td>.7125</td>
</tr>
<tr>
<td>L:H 2</td>
<td>-.360</td>
<td>211</td>
<td>-2.339</td>
<td>.0203</td>
</tr>
<tr>
<td>L:H 3</td>
<td>-.332</td>
<td>35</td>
<td>-2.028</td>
<td>.0502</td>
</tr>
<tr>
<td>L:H 4</td>
<td>.219</td>
<td>91</td>
<td>2.559</td>
<td>.0193</td>
</tr>
<tr>
<td>L:H 5</td>
<td>.235</td>
<td>95</td>
<td>2.131</td>
<td>.0356</td>
</tr>
<tr>
<td>L:H 6</td>
<td>-.190</td>
<td>118</td>
<td>-1.724</td>
<td>.0873</td>
</tr>
</tbody>
</table>

Table 5
But what about how much people use particular applications? Not surprisingly email was a mixed bag. For four of the participants there was no significant difference on email usage for their most productive and least productive time segments. For one of the HCI researchers, a significant difference on the amount of email usage was found on low productivity (M=.537, SD=.505) and high productivity (M=.318, SD=.471), t(83)=2.063, p<.05. For this research using email was associated with not working. For the programmer, a significant difference on the amount of email usage was found on low productivity (M=.231, SD=.430) and high productivity (M=.547, SD=.497), t(118)=-3.208, p<.001. For him email was associated with working.

The two participants, who used spreadsheets in the selected time segments, used them more for work than for work. One of these differences was not significant and the other was marginally significantly: low productivity (M=0, SD=0) and high productivity (M=.172, SD=.378), t(211)=-1.755, p<.08. But for all participants, a significant difference was found on the amount of spreadsheet usage on low on low productivity (M=.017, SD=.131) and high productivity (M=.108, SD=.311), t(643)=-3.739, p<.0002.

Similarly all six participants used word processing and for all but one of these word processing was associated with higher user ratings. Four of these
differences were significant as was the result for the person who used word processing not for work (see Table 6.)

Table 6

<table>
<thead>
<tr>
<th></th>
<th>Mean Diff.</th>
<th>DF</th>
<th>t-Value</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>L:H Total</td>
<td>-0.132</td>
<td>643</td>
<td>-3.789</td>
<td>&lt;.0002</td>
</tr>
<tr>
<td>L:H 1</td>
<td>-0.135</td>
<td>83</td>
<td>-2.158</td>
<td>0.0338</td>
</tr>
<tr>
<td>L:H 2</td>
<td>0.389</td>
<td>211</td>
<td>-3.075</td>
<td>0.0024</td>
</tr>
<tr>
<td>L:H 3</td>
<td>0.199</td>
<td>35</td>
<td>-1.117</td>
<td>0.2716</td>
</tr>
<tr>
<td>L:H 4</td>
<td>0.019</td>
<td>91</td>
<td>-0.849</td>
<td>0.3984</td>
</tr>
<tr>
<td>L:H 5</td>
<td>0.132</td>
<td>95</td>
<td>-1.438</td>
<td>0.1537</td>
</tr>
<tr>
<td>L:H 6</td>
<td>0.154</td>
<td>118</td>
<td>-4.100</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

T-tests on amount of use of different applications by the most and least productive time segments revealed some similar findings. A significant difference on the amount of use of word processing applications was found on the low productivity (M=.098, SD=.298) and high productivity (M=.229, SD=.421), t(643)= -3.789, p < .0002. A significant difference on the amount of use of spreadsheet applications was found on the low productivity (M=.017, SD=.131) and high productivity (M=.108, SD=.311), t(643)= -3.739, p < .000. A significant difference on the amount of use of instant message applications was found on the low productivity (M=.109, SD=.313) and high productivity (M=.008, SD=.092), t(643)= 6.298, p < .0001. Participants used word processing and spreadsheet applications more for work and instant message services when not working. No one used programming and mathematical or statistical applications during the selected time segments and no significant differences were found on web and email usage.
A few interesting trends across participants were found. High user ratings of productivity were correlated with more typing, $r(62) = .313 \ p < .01$ when measured by segment and $r(2747) = .061 \ p < .001$ when measured by switch. So if people were typing a lot, they were more likely to be working. High user ratings measured by segment was also associated with increased switching between applications $r(62) = .264 \ p < .05$. Though this may seem surprising at first, this along with the relationship between high user ratings and typing and findings about application use suggests that people might simply be doing more when they are working than when they are not.

4.1.3.3 Discussion

Switching between applications does not necessarily indicate that someone is not working. After all, there are many important tasks that require multiple applications and a switch between applications does not always indicate a switch in tasks. Across individuals some applications seem to be generally indicative of working: word processing, spreadsheets, and programming environments, for instance. Some applications are clearly used for both work and not work, email and web being the best examples. And some are rarely used for work, such as chat and email. (Though this is not always true as some companies use chat services for internal communication and this author often uses instant messenger to communicate work-related information with colleagues in other countries.)
This method of study seems like an extremely fertile way of determining when people are working and when they are not. Further work may take several directions. Some evidence that is not presented in this paper indicates that time of day may predict amount of work, both on aggregate and individually. Also investigating particular patterns of switches (e.g., email and chat versus email and word processing) might yield a better sense of when someone is productive or not. For instance, typing versus clicking versus viewing in particular applications may indicate the type of task being accomplished. For instance, if I am sitting in front of my word processing program and not typing for a while and then suddenly start to type furiously, I may be first planning and then executing. These are important stages to identify since depending on the task, the user may be more or less willing to be interrupted by different types of information.

4.2 Design Outcomes and Implications of These Findings

4.2.1 Monitoring Application Switching

Taken together, the work habits survey and WorkSniffer study demonstrates that it is possible to understand something about people’s productivity from watching how they switch tasks. Thus, software may be able to identify some features of when people are working and when they are not with some simple algorithms.
The studies demonstrate that in a variety of ways software could likely tell when people are working, when they can comfortably be interrupted and by what. This could be accomplished with minimal invasion of privacy. If the software knows what applications particular people use for work versus not work, they can use that information to determine when people are working and when they are not.

Software could have additional understanding of how applications are used together (e.g., switching rapidly with some control character typing between Microsoft Word and a web browser may indicate cutting and pasting whereas longer switches may indicate losing focus.) This would allow the software to analyze chunks of application switching as single tasks, which would tap in better into how people are actually doing single tasks with multiple applications and be able to model individual action. Whether a user is typing, the category of keys they are pressing (characters, numbers, control, delete, etc.), how much they are clicking and how long they are viewing may indicate different activities with one application. Table 7 shows some possible ways of interpreting different behaviors though these are certainly not definitive answers. Of course, both of these types of measurement would be application specific (e.g., editing in Microsoft Word looks quite different than in emacs.) and would be an imperfect measurement tool, but may suit some application where some error is acceptable.
<table>
<thead>
<tr>
<th></th>
<th>Microsoft Word</th>
<th>Eudora</th>
<th>Netscape</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Typing</strong></td>
<td>If rapid, then generating new text. If slow with scrolling and control and delete keys, then editing.</td>
<td>If rapid, writing email. If slow with scrolling and control &amp; delete key, then editing or organizing.</td>
<td>If few words, searching or, if associated with clicking filling in forms. If many, writing.</td>
</tr>
<tr>
<td><strong>Clicking</strong></td>
<td>Perhaps formatting.</td>
<td>Perhaps searching.</td>
<td>Browsing</td>
</tr>
<tr>
<td><strong>Waiting</strong></td>
<td>Planning (particularly if followed by rapid typing) or not working.</td>
<td>Reading or not attending.</td>
<td>Reading, not attending, or just waiting.</td>
</tr>
<tr>
<td><strong>Rapid Switching</strong></td>
<td>If using control characters, cutting and pasting.</td>
<td>Checking email or if using control characters, then cutting and pasting.</td>
<td>Searching</td>
</tr>
</tbody>
</table>

**Table 7: Possible Interpretations of Different Behavior with Applications**

It is possible to measure lost time from task switching. If, for instance, Jackson et al. (2001) finding that it takes an average of 64 seconds for his participants to recover from email-checking holds true for a larger population, then lost time from switching can also be calculated.

One goal for this project was to limit the invasion of privacy and learn what can be determined from limited information. Certainly software like this would have a better understanding what people are doing if it tracks what people are typing and what content they are viewing. From this information it could better determine when people were doing work-related activities and when a
change in application indicates a change in task. Knowing where people are looking and what else in their environment attracts attention would also be helpful for software and interface designers, but also for users themselves.

4.2.2 Thoughts on Monitoring Tools Used By Organizations and Individuals

There are many reasons to have tools that monitor our work habits and also many reasons to avoid it. The most obvious application of this type of software is to measure white-collar productivity, which has traditionally been a difficult problem. One could see this possibility in one of two ways: as a killer-ap that will allow employers to know when and how much employees are working or as a frightening Big Brother scenario that no ethical developer should attempt. Both interpretations strike this author as valid. Since companies increasingly monitor their employees, it is critical that this monitoring be as productive for all involved as possible and not strip the trust from the relationships between the employer and employee by misbalancing power and control.

A monitoring application is frightening and unethical only if takes power and control away from the employee and gives it to the employer. If one views the employee and employer not as antagonists but rather as two parties with some common interests and some differing goals, the problem changes. Both the employer and the employee want to feel good about the employee’s work.
Transparency and common ground are two critical components of building a healthy organization. Unless the information observed is obviously helpful to the well-being of the organization, monitoring people is not worth the risk of harming the relationship between the employer and employee. If a monitoring system does exist, it is critical that all parties understand its limitations. (For instance, perhaps the system will not identify when an employee is having an important conversation with her manager or know when email from outside the company is work-related.) Second the employee should at a minimum be able to view what the employer sees. Preferably the employee can comment or further exert control over the information. Both parties should recognize that taking breaks and building social relationships within an organization are a necessary component of productive work habits and that it is both the employee and employers responsibility to facilitate work environments that allow for such activities. If this common ground can truly be established, then software that tracks behavior might well be tools for individual and aggregate self-reflection, rather than monitoring.

Organizations should be wary of disrupting their employees through monitoring them since monitoring might come at great cost to the employer. If employees know their switching is being watched, it may be more difficult for the to work productively. Since negative affect has a cost to productivity, simply being upset by monitoring could hurt more than it helps. And, in addition, employees may resent being monitored and, therefore, be less productive on their official
tasks and less motivated to help the company with tasks beyond their duties. Thus it may be more beneficial to the organization for employees to be able to switch as they deem fit and feel positively about their work than to experience monitoring.

Organizations can contemplate productivity in other less intrusive ways of understanding and encouraging desired work habits. They may present the literature described in this paper to their employees and suggest from there that they all be more careful about how they allow themselves to be interrupted. They could support controlled organization-wide discussions of the corporate culture and how to support both space for work and social interactions. What is the social structure? Who is interruptible and under what conditions? How quickly should people to others calling, sending email, or stopping by their office? How do people take breaks? When do people get frustrated by interruptions? How do different individuals work habits support different social interactions?

Organizations that use email may consider examining the types of email clients and other push technologies that their employees use. They might create policies of when and how to interrupt. Some strategies for email that might help include 1.) cc:ing emails only when necessary 2.) limiting jokes to the beginning and end of days 3.) maintaining a separate account for personal email and 4.) writing succinct emails in which the desired response is
immediately clear. (Jackson et al found that when employees were limited to a new Short Message System email system, 80% reported a time savings (2001).) If people check their email constantly but feel most email is non-urgent, perhaps it would be helpful for them to restrict their email checking to once an hour or several times a day. Perhaps certain web sites that are never used for work should be restricted while someone prefers to be working. (Having recently purchased a house, controls on ebay and craigslist would have been helpful for me while writing this thesis!) Finally they might encourage their employees to use a self-monitoring tool like the one described in the following section, but be clear that this information is for the eyes of that particular employee only. Such suggestions must be implemented with care as they may involve introducing additional requirements for employees that are not related to their core tasks. The goal is not to create more work but to best support a range of productive work styles.

4.2.3 One Prototype Based on these Findings

Individuals, too, are likely to be interested in how they work. If many of us find that it is difficult to keep track of where our time goes during the day, software that transparently monitors and displays what we are doing may provide a unique tool for self-reflection. For this software that reflects our work habits back to us to be successful, it should be immediately relevant without being distracting, while being completely transparent about what it
tracks and displays. This is a very delicate balance to achieve. Less may be more.

*SimplyFocus* is a task interference self-monitoring tool to assist users with maintaining deep focus and taking breaks in a timely manner. It is not presented as the ultimate solution, rather as an implementation of some of the ideas culled from the previously discussed investigations. The purpose of this application is to help users to self reflect upon their work.

*SimplyFocus* is, by design, a very simple system. Minimal tools like this one may help users to maintain their focus and not to distract them when they didn’t wish to be distracted. It provides real-time feedback which may be important to keep the self-reflection going but it does so gently and allows the user to turn this feature off if she likes. It also presents accumulated displays to contemplate periodically, perhaps as a work break.

Tools such as this one should be clear about what information is collected and where it is going. In this case, it is important to demonstrate that the feedback is going primarily to the user a (as opposed to the software or user’s employer) in order to help the him to learn about his individual work style. The application was designed to encourage the user to think about their work habits, not to observe and judge what they do. As a result the system only bases its judgments upon whether someone is working on a simple set of
algorithms that take into consideration what application they are using, how productive they label that application as being, and how much they are typing in the application.

The system is built upon *WorkSniffer* and records the exact same information that the previous software did: when people switched applications, what applications they used and how much they typed alphabet characters, numeric characters, shift, control, up, down, left, and right.

This feedback from the application can be viewed in two ways: either reflecting the current status or work or aggregated over time. A menu that can be accessed either through the taskbar or graph version of the application allows for customization. Here they can select which task-bar icons they wish to see, whether they want to see the aggregated data graph, and whether they want the graph to always be on top. They can also indicate how much they use different categories of applications (writing, office, programming, email, entertainment, telnet/ssh, and web) on a 5 point scale (most productive, usually productive, sometimes productive, a little productive, and unproductive.)

![Figure 10: Taskbar View](image)
application viewing option is a real-time taskbar icon that changes how it looks as the user works. As the user switches from applications that they describe as more or less productive, the icon changes. This taskbar icon can be customized to look one of three ways: a circle that changes color, a face that smiles and frowns, and as a diamond that pulses more rapidly to indicate more work.

Users can also see their work aggregated over time by double-clicking on the taskbar icon. Then a graph appears that illustrate which applications were open by color and how much they are used by height along the y-axis. The x-axis represents time.

In informal discussions and presentations, reactions to this software application were polarized. Some individuals were appalled that a computer would observe their work habits. Others were incredibly positive and welcomed such an interface. These users felt sure that they did not self monitor well, but weren’t sure exactly what they did less than optimally or what efforts they could make to improve. They felt that this type of application was exactly what they need.

4.2.4 Specific Design Ideas for Systems that Support Self Aware Work Habits

The implementation of SimplyFocus could be improved in several ways. The end of the workday likely is a good time to encourage self-reflection. Software like SimplyFocus could either rely on the users to intuit what they felt they accomplished themselves by presenting statistics about the day or assess
productivity from the data and present its evaluation. The user could decide which version he prefers. Either way, at the beginning of the day, the user could indicate his goals (e.g., writing a document, editing it, coding part of a program.) At the end of the day, if he prefers to evaluate himself then the software could present 1.) how much he looked and typed in particular applications that are productive and that are unproductive for the defined tasks 2.) when the user typically used which applications and how long he used them. (Does my 10 minutes of email in the morning actually take an hour? Though I feel I write best in the morning, do I procrastinate and type frantically around 6:30 PM most days?) and 3.) how much he switched to self-defined distractions such as email or chat while working with productivity applications. If the software was to interpret productivity, the sign off page may indicate the same type of information but do so based on algorithms based on the previous findings. At the end of the day, based on the applications used when, how much and how they were used (e.g., typing versus viewing) the software could indicate the progress towards the defined goals. This time of self-reflection would also be a good point for the software to ask for feedback on how it is doing: both on accuracy of assessment (It’s true that I didn’t accomplish my goals ) and relevancy (but my goals changed when I noticed a paper deadline is today. ) The user could also choose to change the defined goals during the day to make the software more accurate. Eventually the software should be able to understand how goals change throughout the day.
Also *SimplyFocus* could be extended to turn off certain applications, or at least their notification ability, off. This way, users could avoid undesirable interruptions. A system like this may allow users to either 1.) choose work intervals during which time they will be left alone (e.g., from 9 AM to 11 AM when I have noticed from my Graphing View of *SimplyFocus* I am most likely to word process) or 2.) choose applications or activities during which they should not be interrupted (e.g. whenever I am using my word processor or statistics application.)

4.2.5 *Interface Ideas for Email*

If users are not good at managing their email in an efficient manner, then interface designers may wish to create ways to limit the distraction of email arrivals, provide better tools for visualizing information, and interrupt more appropriately based on the current task and the importance of the incoming mail. Particular interfaces that more obviously push new email at the user may encourage more inefficient switching. If this is the case (and it is an open research question whether it is,) designers might consider several design options.

Alerts can be context specific. If I am surfing the web at work, an aggressive loud noise maybe appropriate for work-related email. But if I am at home and a work email arrives, perhaps I want to be notified more subtly or even *not* notified unless the message is critical.
Another option would be to have different methods of sorting email or other incoming messages. Already mail clients support labeling the importance of email. Though this feature is often misused and therefore rendered useless, the idea could be repurposed. Why not be able to sort email by a variety of features other than timestamp, sender and receiver? Rating messages on urgency versus importance or urgent for me versus urgent for you might help people better describe when a response is critical. Setting timeframes for suggested response times (honestly) might help people manage their time efficiently and remember when they need to work on a particular task. Sorting by project or topic would provide a platform for better filtering and allow users to know whether an email is related to the task at hand.

People use email for more than raw communication. For many it is a contact list, a calendar, and a task list. More email applications could mark particular emails as tasks or events or both as Microsoft Outlook does.

4.2.6 General Implications for User Interface Designers

Whether designing an email client, an agent that manages incoming information, or a productivity tool, designers need to understand not only the limits of their users’ attention, but also what the users will realistically be doing, given many possible pursuits fighting for their attention. This does not mean that a particular interface should be aggressive so as to win the
attention-getting war. Rather it should be built with a sensitivity to our relationship with our work.

Designers may look to the cognitive science and HCI literatures to know and how efficient work can be supported for different users (personal styles, novice versus expert, age differences, etc..) They must know what features of their interface support different work styles and how the interface distracts people from their work and impede quality output. Will their users be easily distracted by email? Will they be willing to learn complicated sets of keystrokes so that later they will be facile and quick with an interface?

We may not keep our long-term goals in mind moment-to-moment, but we would like it if our interfaces did or at least helped us to do so. And we certainly do not want to waste our limited working memory on push interfaces that distract us from our primary goals. Since users may feel compelled to check the contents of disrupting tasks regardless of importance and since they may actually perform less well when they complete the distracting task, designers of push or disruptive technologies should take great care in considering the mechanism by which the interruption occurs and how important the primary task is. Some users (this one included) may appreciate feedback about how well we are managing our time while others may not.
We can determine when and how to interrupt users by comparing the importance and content of primary and secondary tasks and their tolerance to errors, e.g., a spelling error in an email to a friend is tolerable but one in a thesis abstract is not. Designers could hide low-importance or currently irrelevant distractions from the user. Or they could interrupt more when the user is engaged in low-importance work. McFarland (1998) found that his participants did well when they could negotiate when an interruption occurred, but this also meant that they might switch right away and, similar to participants in this study, performed less well on the interrupting tasks. Still, it may be beneficial to give users the option to switch, or, perhaps even better, to have the system turn notification off when the primary task work is critical. Designers whose application allow interruptions from other applications (such as an instant message application that informs users when they have email) should consider whether they want their users to switch applications and whether they are concerned about how their users perform on their secondary tasks.

Customizable alert systems for email and other applications would allow users to create generic rules about which messages should interrupt them. A model of tasks within applications could be classified as working or non-working. If I am emailing my husband, I am doing personal activities. If I am emailing my advisor, I am working. And if I am looking at dictionary.com, I am working but if I am looking at cnn.com, I am surfing. Or they may be more sophisticated: if I am emailing my husband and there is code in the body of
the message, I am working or if I am using google, but searching for information about gardening, I am not working. If I am searching for research papers, I am working. Even better would be more specific rules that consider what the current task is and how the content of the email may be matched to it. If I am deep in thought and in Miyaka and Norman’s evaluation stage or if the email content is poorly matched to my current task, I shouldn’t be interrupted. But if the content is well-matched and I am in the planning stage, then perhaps I should be made aware of the message.

This matching could also work in the reverse direction. Suppose I am stuck and don’t know what to do next. The system could scan for people I’ve emailed about related issues and have that page in front the next time I look at my email.

These findings could be applied not only to the content and functionality of an interface, but also to the interface itself. A poorly designed interface can divide our attention as much as multiple tasks. The version of Microsoft Word I use often highlights words as misspellings that are correctly spelled. I find that seeing the words underlined in red divides my attention and I often stop writing to right click on the word so that highlighting goes away. Problems such as this one come at a real cost to the user. In my case, I am commonly interrupted while planning what to write next and cannot remember what I was thinking when I return to my primary task.
Designers may already consider how many clicks or keystrokes it takes to complete a task. But they may also wish to consider how much thinking needs to occur in order to remember what actions to take. High-load interfaces are distracting and likely take more time without providing added benefit to the user. Of course the same interface that loads a novice user could be faster for an expert and this too should be considered.

Further investigation into individual management styles of work (such as future investigations similar to the WorkSniffer study) may lead to improved understanding of how to build customizable or adaptive interfaces. It may be the case that users have individual and importantly variant ways of managing and distributing these kinds of interactions. In this case, designers should likely provide means of supporting diverse interaction styles.

5 Conclusions

Computers are good tools for working on an ever-increasing number of tasks, such as word processing, programming, and communicating. The drawback of this flexibility is that distraction is always a click away. As a result, we may switch between tasks inefficiently. Based on previous work on attention, we know there is often a performance time and quality cost to switching tasks. But less is know about the choices that people make when switching between tasks
with computers and whether the habits that evolve from these choices are productive.

From this work we've learned that people do not think about long-term efficiency when they work, at least not with their email. Certain computer applications are associated as work-related for many users and by watching application usage, we know something about whether users are productive. Profiles of individual work habits can likely be built based on very basic information such as application switching and amount of typing and clicking. Rapid switching sometimes indicated low productivity, but sometimes can also mean high productivity. Future work into how people use multiple applications to accomplish single tasks would better define how this profiling should occur and when switching indicates high and low productivity.

I presented an application for individuals to self-reflect upon their own work habits. Software like this should be immediately relevant without being distracting, while being completely transparent about what it tracks and displays. Users may be able to help themselves by reflecting on their own work. The hope is that by giving people some limited feedback about how they are working, users will be able to focus more when they choose to and less when they need to take a break. This is a very delicate balance to achieve. Less may be more.
I suggested that many interfaces may unintentionally or intentionally interrupt or distract users from both their long-term and short-term goals. This can happen through untimely pushes of information, ease of switching to a more entertaining, but less fulfilling task, or plain old confusing interfaces. Interface designers may be able to support some delicate combination of the user's short-term and long-term goals (which may conflict) by considering not only the limits of our attention but also how and why we may choose to allow ourselves to be distracted.

Each individual works with her computer and with the world outside it in a unique way. How she wishes to work can change from day to day and hour to hour, sometimes in ways that help her productivity and fulfillment and sometimes not. The goal in observing people's work habits should not be to restrict their usage and therefore the expressions of their personal style, nor should it be to know more about them than they do. The ultimate hope is that our machines will understand our wishes and will be able to support our long-term goals of fulfilling and productive work, while understanding we do not walk a straight path along the way.
6 References


Humphreys, M. Proactive interference and complexity; Journal of Experimental Psychology, Learning Memory and Cognition, Washington; May 2001; Vol. 27, Iss. 3; pg. 872.


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


O’Conaill, B. and Frohlich, D. Timespace In The Workplace: Dealing With Interruptions . CHI 95.


Rehder, B. Interference between cognitive skills; Journal of Experimental Psychology, Learning Memory and Cognition, Washington; Mar 2001; Vol. 27, Iss. 2; pg. 451.


Ruthruff, E. Why practice reduces dual-task interference; *Journal of Experimental Psychology, Human Perception*. Washington; Feb 2001; Vol. 27, Iss. 1; pg. 3.


7 Appendixes

7.1 Task-switching Study Instructions

*Condition 1: Seen experiment condition.*

In this study you will solve puzzles. When you start up the application, you will see five blue dots on the left side of the screen. Each dot corresponds with a puzzle. If you click on a dot, you'll get a puzzle. You can do the puzzles in any order you want. You can go back and forth between puzzles. But once you hit submit for a puzzle, the dot will go away and you won't be able to go back to it. While you are working red dots will be appearing on the right side of the screen. There will be 10 red dots in total. Each red dot corresponds with a puzzle. Just as before, if you click on a dot, you get a puzzle. You can do the puzzles in any order you want. You can go back and forth between puzzles. But once you hit submit for a puzzle, the dot will go away and you won't be able to go back to it. The payment scheme is as followed. If you solve a blue dot puzzle perfectly, you will earn one dollar. Solving it perfectly means you make the numbers add up to 100. If you make the numbers add up to any number between 80 and 120, except for 100, you will make 25 cents instead. For the red dot questions, you will make on average 25 cents. Sometimes you will make more. Sometimes you will make less, but on average you will make 25 cents. You will see the value of each red dot question when you click on it. You will have 20 minutes to complete the study. I will stop you after 20 minutes. Any questions so far? Ok, now I would like to show you an example of what the blue dot questions and red dot questions will look like. [Show the example.] Any questions? Ok, please start.
Condition 2: Unseen experimental condition.

In this study you will solve puzzles. When you start up the application, you will see five blue dots on the left side of the screen. Each dot corresponds with a puzzle. If you click on a dot, you'll get a puzzle. You can do the puzzles in any order you want. You can go back and forth between puzzles. But once you hit submit for a puzzle, the dot will go away and you won't be able to go back to it. While you are working red dots will be appearing on the right side of the screen. There will be 10 red dots in total. Each red dot corresponds with a puzzle. But the red dots will be hidden. In order to see the red dots, you will need to click on the Check for other questions button to reveal them. Just as before, if you click on a dot, you get a puzzle. You can do the puzzles in any order you want. You can go back and forth between puzzles. But once you hit submit for a puzzle, the dot will go away and you won't be able to go back to it.

The payment scheme is as followed. If you solve a blue dot puzzle perfectly, you will earn one dollar. Solving it perfectly means you make the numbers add up to 100. If you make the numbers add up to any number between 80 and 120, except for 100, you will make 25 cents instead. For the red dot questions, you will make on average 25 cents. Sometimes you will make more. Sometimes you will make less, but on average you will make 25 cents. You will see the value of each red dot question when you click on it. You will have 20 minutes to complete the study. I will stop you after 20 minutes. Any questions so far? Ok, now I would like to show you an example of what the blue dot questions and red dot questions will look like. [Show the example.] Any questions? Ok, please start.

Condition 3: Add up to 100 first control condition.

In this study you will solve puzzles. When you start up the application, you will see five blue dots on the left side of the screen. Each dot corresponds with a puzzle. If you click on a dot, you'll get a puzzle. You can do the puzzles in any order you want. You can go back and forth between puzzles. But once you hit submit for a puzzle, the dot will go away and you won't be able to go back to it. After you solve the five blue dot questions, 10 red dots will appear on
the right side of the screen corresponding with 10 more puzzles. Just as before, if you click on a dot, you get a puzzle. You can do the puzzles in any order you want. You can go back and forth between puzzles. But once you hit submit for a puzzle, the dot will go away and you won't be able to go back to it. The payment scheme is as followed. If you solve a blue dot puzzle perfectly, you will earn one dollar. Solving it perfectly means you make the numbers add up to 100. If you make the numbers add up to any number between 80 and 120, except for 100, you will make 25 cents instead. For the red dot questions, you will make on average 25 cents. Sometimes you will make more. Sometimes you will make less, but on average you will make 25 cents. You will see the value of each red dot question when you click on it. You will have 20 minutes to complete the study. I will stop you after 20 minutes. Any questions so far?

Ok, now I would like to show you an example of what the blue dot questions and red dot questions will look like. [Show the example.] Any questions? Ok, please start.

**Condition 4: Sequence completion first control condition.**

In this study you will solve puzzles. When you start up the application, you will see 10 red dots on the right side of the screen. Each dot corresponds with a puzzle. If you click on a dot, you will get a puzzle. You can do the puzzles in any order you want. You can go back and forth between puzzles. But once you hit submit for a puzzle, the dot will go away and you won't be able to go back to it. After you solve the 10 red dot questions, five blue dots will appear on the right side of the screen corresponding with five more puzzles. Just as before, if you click on a dot, you get a puzzle. You can do the puzzles in any order you want. You can go back and forth between puzzles. But once you hit submit for a puzzle, the dot will go away and you won't be able to go back to it. The payment scheme is as followed. If you solve a blue dot puzzle perfectly, you will earn one dollar. Solving it perfectly means you make the numbers add up to 100. If you make the numbers add up to any number between 80 and 120, except for 100, you will make 25 cents instead. For the red dot questions, you will make on average 25 cents. Sometimes you will make more. Sometimes you will make less, but on average you will make 25 cents. You will
see the value of each red dot question when you click on it. You will have 20 minutes to complete the study. I will stop you after 20 minutes. Any questions so far? Ok, now I would like to show you an example of what the blue dot questions and red dot questions will look like. [Show the example.] Any questions? Ok, please start.
7.2 Email Survey Questions

How many emails do you get a day?

- <10
- 10-25
- 25-50
- 50-100
- 100-500
- >500

How much time do you spend on email a day?

- 0-5 minutes
- 5-15 minutes
- 15-30 minutes
- 30 minutes to 1 hour
- 1 to 2 hours
- 2 to 3 hours
- over 3 hours

Is your software set up to automatically check for new email?

- Yes
- No

a. If you answered yes to the previous question, how often does your software check?

- Every 0-5 minutes
- Every 5-15 minutes
- Every 15-30 minutes
- Every 30 minutes to 1 hour
- Every 1 to 2 hours
- Every 2 to 3 hours
- Over every 3 hours

b. If you answered no, how often do you check to see if new email has arrived?

- Every 0-5 minutes
- Every 5-15 minutes
- Every 15-30 minutes
- Every 30 minutes to 1 hour
- Every 1 to 2 hours
- Every 2 to 3 hours
- Over every 3 hours

How often do you read the email you receive? (We are interested in how often you read the body of the message, not just looking at who the mail is from and what the subject line is.)

- Every 0-5 minutes
- Every 5-15 minutes
- Every 15-30 minutes
- Every 30 minutes to 1 hour
- Every 1 to 2 hours
- Every 2 to 3 hours
- Over every 3 hours

How often do you respond to email?

- Every 0-5 minutes
- Every 5-15 minutes
- Every 15-30 minutes
- Every 30 minutes to 1 hour
- Every 1 to 2 hours
- Every 2 to 3 hours
- Over every 3 hours

What percentage of your email requires immediate attention?

- 5%
- 10%
- 25%
- 50%
- 75%
- 100%
What is the most efficient time schedule to read your email?

<table>
<thead>
<tr>
<th>Every 0-5 minutes</th>
<th>Every 5-15 minutes</th>
<th>Every 15-30 minutes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Every 30 minutes to 1 hour</td>
<td>Every 1 to 2 hours</td>
<td>Every 2 to 3 hours</td>
</tr>
</tbody>
</table>

If you don’t check your email in the most efficient way, why is that?
7.3 Work Habits Survey Questions

How old are you?

What is your occupation? Highs chool student, undergraduate student, graduate student, academics, business, retired, other.

If you are a student or if you work, how many hours a day do you work?

How many hours a day do you use a computer for work-related activities?

How many hours a day do you use a computer for non work-related activities?

When working intensely on a project without distractions, which applications do you commonly have open?

- word processing
- email
- spreadsheet
- web editor
- programming environment
- web browser
- statistics/math
- organizer
- games
- graphics/audio/video editing
- music listening

Please list any other applications.

When working intensely on a project without distractions, please rank in order the applications you use the most in the order that you use them. (1= most used, 2= second most used, etc.)

When working on a project but not working very hard, which applications do you commonly have open?

- word processing
- email
- spreadsheet
- web editor
- programming environment
- web browser
- statistics/math
- organizer
- games
- graphics/audio/video editing
- music listening

Please list any other applications.
When working on a project but not working very hard, please rank in order the applications you use the most in the order that you use them. (1= most used, 2= second most used, etc.)

When using a computer not for work, which applications do you commonly have open?

word processing  email  spreadsheet  web editor programming environment  web browser  statistics/math  organizer  games  graphics/audio/video editing  music listening

Please list any other applications.

When using a computer not for work, please rank in order the applications you use the most in the order that you use them. (1= most used, 2= second most used, etc.)
7.4 Excerpt of an Example WorkSniffer Log File

01/09/2003 15:12:45 Monitoring Started
User: cM

01/09/2003 15:12:50 Switched to Windows Window [0 0 0 0 0 0]
01/09/2003 15:12:53 Switched to Productivity [0 0 0 0 0 0]
EOF

01/09/2003 15:28:57 Monitoring Started
User: cM

01/09/2003 15:29:08 Switched to Windows Window [0 0 0 0 0 0]
01/09/2003 15:29:09 Switched to Productivity [0 0 0 0 0 0]
01/09/2003 15:29:12 Switched to Microsoft Visual C++ [0 0 0 0 1 0]
01/09/2003 15:29:17 Switched to Productivity [0 0 0 0 0 0]
01/09/2003 15:29:21 Switched to Windows Message Box [0 0 0 0 0 0]
01/09/2003 15:29:22 Switched to Productivity [0 0 0 0 0 0]
01/09/2003 15:29:25 Switched to Windows Message Box [0 0 0 0 0 0]
01/09/2003 15:29:27 Switched to Productivity [0 0 0 0 0 0]
01/09/2003 15:29:31 Switched to Microsoft Visual C++ [0 0 0 0 0 0]
01/09/2003 15:29:33 Switched to Productivity [0 0 0 1 0]
EOF

01/09/2003 16:36:31 Monitoring Started
User: cM

01/09/2003 16:37:01 Switched to Windows Window [0 0 0 0 0 0]
01/09/2003 16:37:02 Switched to Internet Explorer [0 0 0 0 0 0]
01/09/2003 16:37:03 Switched to Productivity [0 0 0 0 0 0]
01/09/2003 16:37:04 Switched to Internet Explorer [0 0 0 0 1 0]
01/09/2003 16:37:10 Switched to Productivity [0 0 0 1 0]
EOF
7.5 Essay Experiment Instructions

In this experiment, you will be asked to write an essay. The essay should be whatever length you feel is appropriate. You will have 10 minutes to write this essay. You can use any application on this computer to research this essay, including any web sites or digital libraries that might help you. Please take as much of the 10 minutes as you need but if you happen to finish early, you may do whatever you like.

To begin this experiment, please

1.) Doubleclick the  ClickMe  icon in this folder.
2.) Open the  Essay  file located in this folder.
3.) In the  Start Time  box on your essay file, type the start time listed on your computer.
4.) When you are done, type the end time in the  End Time  box and be sure to save your file.

[Then either the participants reads]
The topic of your essay is your favorite food or beverage. Please describe your favorite food or beverage, what you like about it, why you like it and, if applicable, what memories or associations you have with this food.

[or the participant reads]

The topic of your essay is why is it ok to push a fat man off a bridge if it would prevent a bridge with 500 people on it from falling. You should assume that you are the person who will push the fat man off of the bridge. Please describe why you would think this was the right thing to do and how you would do it.
7.6 Work Logging Observation Instructions

What I’d like you to do is install this software, pick a day to when you will be using your computer and let me know the night before or in the morning that you will run it that day. Then leave it running on your computer all day. About an hour before you quit for the day, email me the file. I’ll then ask you to fill out a very brief questionnaire and will go through the log file with you for a few minutes to find out what you were doing during different periods of time that day. The reason I want you to send me the file before you quit for the day is so that we will have a chance to talk about the log file. This could be in person or by phone, whichever is easier for you.

That’s it. Feel free to ask questions if something doesn’t make sense or if you would like to do things slightly differently.

Here’s the installation instructions:

1.) Create a folder in your C: directory. Extract all of the files from the zip file into this directory.

2.) Create a second folder in your C: directory called 'Data.' Leave this folder empty for now.

3.) Create a shortcut to the file, 'ClickMe.exe' that is now located in the first directory you created. Put the shortcut in your startup directory. Your startup directory is probably something like: C:\Documents and Settings\Administrator\Start Menu\Programs\Startup.

4.) Restart your computer. If you did everything correctly, three files will be created in your newly created data directory.