9-10-2009

Time delays and system response times in human-computer interaction

Noah Stupak

Rochester Institute of Technology

Department of Psychology

College of Liberal Arts

TIME DELAYS AND SYSTEM RESPONSE TIMES
IN HUMAN-COMPUTER INTERACTION

A Thesis in

Applied Experimental and Engineering Psychology

by

Noah Stupak

Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Science

September 10, 2009

Esa M. Rantanen, Ph.D.

Associate Professor of Psychology

Thesis Adviser and Chair of the Thesis Committee

Nicholas DiFonzo

Professor of Psychology

Reader

Andrew Herbert

Associate Professor and Chair, Department of Psychology,

Reader

Abstract

This study sought to determine a type of graphical representation of system response time that would be most beneficial to the user in terms of task performance. Specifically, I examined which type of progress bar would allow the user to return to working with the system most efficiently while performing other concurrent tasks, and how well the user performed these tasks. The different types of progress bars studied included segmented and continuous progress presentations, and linear, accelerating, and decelerating progress behaviors. The results indicate that different representations of system response time affected performance on the two tasks, with the continuous progress bar resulting in the best performance out of the two progress bar types and the linear behavior resulting in the best performance out of the behaviors. The results also show that different progress bar speeds and types should be used depending on the desired effect of the progress bar on users.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I owe my sincere gratitude to my thesis advisor Dr. Esa Rantanen, who without his support this project and my degree would not have been completed. His assistance, encouragement, and advice were invaluable to me. I am also greatly indebted to the members of my thesis committee, Dr. Nicholas DiFonzo and Dr. Andrew Herbert, for their help and guidance throughout the thesis process as well as my undergraduate and graduate career at RIT.

I would also like to express my heartfelt gratitude to Melissa Farnand for her endless encouragement and enthusiasm.

Chapter 1: Introduction

The time delay between a user's initiation of a command on a computer and the system's completion and display of the result is known as the system response time (SRT) (Miller, 1968). These times can range from milliseconds to minutes depending on numerous variables, including the complexity of the task and the processing power of the system. SRT is also a major source of frustration for computer users and it can severely degrade the systems' and different applications' usability. Therefore, this temporal aspect of human-computer interaction (HCI) is an important component of a system's overall design.

One would assume that as technology becomes more advanced and processing speeds become faster these delays would lessen. This would be true only if the complexity of the tasks and richness of the content did not increase, too. However, the opposite is true; programs and websites have become increasingly complex offering more and more features, thereby increasing the amount of information that needs to be processed. The more data that are processed, the longer the delays in the system are to be expected.

System Response Times

Time-shared computer systems are one of the earliest catalysts for SRT research (Carbonell, Elkind, & Nickerson, 1968; Morfield, Wiesen, Grossberg, & Yntema, 1969; Miller, 1968). In the past, computer systems were large and expensive; it was not cost effective for a single user to have exclusive access to the system. To keep cost down, multiple users accessed a central mainframe from terminal locations, sharing time on the mainframe. The idle time of one user would allow another to be actively accessing the system. The tradeoff was that users would experience SRT while the central mainframe processed their request. As the cost of computer systems decreased, the use of time-sharing of a central system faded.

Currently, web applications and the Internet have brought a time-sharing model of interaction back into practice. Web application use is becoming increasingly more common, with a third of computer users using at least one web application on a regular basis and more than 50% of students using at least one web application (Mace, 2007). Of this latter group, 49% of application use is solely with web applications and not desktop or client-side applications. Web applications have many advantages over client-side applications. They typically require little or no disk space since they run from a network location, upgrade automatically with new features as they are implemented on the server, and integrate easily with other web applications like e-mail or calendars. Additional advantages of web applications are that they allow the user to work from any location because all data are stored on the server, and they provide cross-platform compatibility as they operate within the web browser window.

Despite these advantages, web applications suffer from the same drawbacks as the first time-shared systems. They rely on files stored on remote servers accessed through the Internet. Therefore, if the connection is interrupted, the application is no longer usable. Under heavy user traffic, the user could experience delays when running the application. Although now a fairly accepted hindrance, the most unpleasant aspect of a website is the SRT duration while waiting for the website to load or respond (Lightner, Bose, & Salvendy, 1996).

SRT and Human Experience and Performance

Little quantitative research exists on graphical SRT representations. Prior studies have focused on looking at a single graphic SRT representation, without comparing how different representations affect users (Kuhmann, Boucsein, Schaefer, & Alexander, 1987; Myers, 1985). The dependent measures in these experiments were common parameters of task performance such as reaction time, number of cursor movements per item, failure rate, and time to complete each item. Additionally, psychophysiological measures such as electrodermal activity, heart rate, and blood pressure were measured as well as subjective reports of mood and discomfort. These studies show that SRT does affect a user's stress level and performance, but because there were no comparisons between graphic representations, it cannot be determined to what extent the graphic relieved the stress and aided performance. In addition, nonlinear graphical representations have had minimal research (Harrison, Amento, Kuznetsov, & Bell, 2007). Harrison's study expanded on previous research on nonlinear representations of SRT, while also seeking to integrate the different representations of SRT and study how these representations influence users.

Initial research on SRT delays and user performance showed that the most important factor in a workplace computer interface was the duration of the SRT, where shortest SRT duration being was the best (Dannenbring, 1983; Martin & Corl, 1986). Later work focused on the predictability of the system delay, and these results indicated that the two most important factors include: the ability to predict the duration of the delay in relation to the type of user action, and the feedback during the delay (Rushinek & Rushinek, 1986). Further research showed that SRT delay, when unavoidable, should be predictable for the user (Shneiderman, 1987). When SRT are constant, the worker can predict when the next step will occur and plan workflow accordingly.

The psychological and physiological ramifications of SRT have been clearly defined by previous research. Studies on the effect of SRT on physiological stress reactions, subjective ratings and task performance have shown that users under time pressure experiencing long SRT displayed negative emotional states, but higher levels of performance (Kuhmann et al., 1987). Kuhmann, Schaefer, and Boucsein's (1990) participants performed a detection and correction task at computer terminal in six trials of 20 minutes each, and the participants who experienced longer SRT had increased skin conductance responses, whereas participants who experienced short SRT had increased blood pressure. In Kuhmann's study there was no feedback to the user during the SRT, so participants were not aware of the duration or expectation of the SRT. Performance on the trials was measured by error rate and work speed. When no time pressure was present, SRT had no effect on physiological responses. Longer SRT without time pressure also led to a decrease in task performance, but a higher subjective evaluation (Kuhmann, Schaefer, & Boucsein, 1990).

SRT has been shown to be one of the strongest stressors in human-computer interaction (Shneiderman, 1987). The stress-inducing factors of SRT were determined to be duration, variability, and expectation (Boucsein, Baltissen, & Euler, 1984). Duration is the amount of time the SRT takes, variability is the variance across a set of SRT, expectation is defined as the user's ability to assess and determine the duration of the SRT, which is essentially an effect of the information provided about the duration of the SRT. Of these three factors, the only one that can be accurately controlled and symbolized for a singular SRT is expectation. If the user is given proper information on the duration of the delay, then the stress effect can be lessened through proper expectation. If SRT variability is mitigated, users report a more positive well-being (Kuhmann et al., 1990).

SRT duration has a direct effect on efficiency and task completion as well as on user frustration. It has been shown that SRT has a significant effect on how quickly a user can complete tasks, in addition to affecting frustration levels of users and having a marginal effect on their efficiency (Selvidge, Chaparro, & Bender, 2001). Another effect of SRT on users is that perceived workload increases as the delay increases (Barron et al., 2004).

A recurring finding in the time estimation and human-computer interaction literature is that because SRTs are unavoidable, they should be predictable for the user. When expectation of the SRT is not possible, users are not able to anticipate the start of the next work step; this is known as temporal uncertainty (Schaefer, 1990). The source of the uncertainty is a lack of information being provided about the duration of the SRT, or improper expectation of the SRT.

Prior research has not looked at how the different representations of SRT affect the ability to engage in multiple complex tasks simultaneously. Studies that have looked at SRT and interaction most commonly used simple tasks such as locating a target among distracters and highlighting it (Schaefer, 1990; Thum, Boucsein, Kuhman, & Ray, 1995), or no specific task at all (Myer & Hildebrandt, 2002). Often, these studies follow a dual-task paradigm, where both tasks can be seen and worked on at the same time. The present study used a task-switching paradigm with the two tasks being unrelated which has advantages. When changing tasks, a "task-set reconfiguration" (Monsell, 2003) is necessary before then new task can be processed and started. The two tasks used in this study are unique enough to not require a drastic task-set reconfiguration, which minimizes the task switching cost and aids good performance. The present study sought to investigate this previously unexamined area.

Graphical Displays of SRT

Due to the many negative consequences of temporal uncertainty, a crucial component of user interface design should be dedicated to creating a representation of SRT that provides information that allows the user to quickly and accurately assess and predict its duration. This temporal display can also be seen as an interface's time affordance. An example of a time affordance on PCs is the hourglass that appears next to, or replaces, the cursor when an operation is being completed. Progress indicators are one of the most common SRT representations. The

Apple Human Interface Guidelines (2006) defined three types of progress indicators to be used with the OS X operating system: (1) the determinate progress bar—displays a thermometer-like bar where the “fill” moves from left to right and should fill in completely before it is dismissed; (2) the indeterminate progress bar—displays a spinning striped cylinder to indicate an ongoing process; and (3) the asynchronous progress indicator—displays a spinning disk, usually where the cursor is located. Note that the indeterminate progress bar is just an asynchronous progress indicator masquerading as a progress indicator.

A good time affordance communicates multiple things to the user (Conn, 1995): (1) acceptance, whether the task has been accepted by the system; (2) scope, the size of the task and duration of time required to complete; (3) initiation, an indication that the task has begun; (4) progress, the rate at which the task is completing; (5) heartbeat, a quick visual indication that the task is still working and has not stopped responding; (6) exception, a notice that the working task requires user input; (7) remainder, an indication of how much of the task remains and/or how much time is left before completion; and (8) completion, an indication that the task has finished, whether successfully or unsuccessfully. These eight items provide information to the user before, during the SRT, and after it has completed. Feedback during the SRT is one of the most important factors for reducing stress and increasing productivity.

Based on these criteria, the Windows hourglass clearly is not a good time affordance because it contains few of these affordances. The Apple asynchronous progress indicator is identical to the Microsoft Windows hourglass in terms of information conveyed. These indicators provide feedback in the form of acceptance (the display of the indicator), initiation (the animation of the indicator), and completion (the removal of the indicator and return of the cursor). It would seem that it also provides a heartbeat (indication that the system is functioning properly), but often the indicator remains even if the computer has stopped responding. Windows also utilizes indicators similar to the progress indicators of Apple. Both determinate and indeterminate bars provide feedback in the form of command acceptance, initiation and completion, with the determinate progress bar also providing information on progress, remaining time, and often the SRT's scope.

Similar to the asynchronous progress indicator is a static or dynamic "wait" message, as both indicate that the system is busy with an indeterminate end. When using an asynchronous progress indicator, two issues are important: first, that the display chosen makes the duration of the waiting period appear minimal. SRTs affect user satisfaction, therefore the user should believe the duration is minimal (Schleifer & Amick, 1989). However, unreliable indictors may diminish any user satisfaction gained from perceived short SRTs. Second; the display should be preferable to a majority of the users, which could mean that the display can be appealing or entertaining. When comparing determinate, indeterminate, and asynchronous displays, it was found that asynchronous and indeterminate displays led users to make longer duration estimates (Meyer, Shinar, Bitan, & Leiser, 1996).

One of the first experiments on the temporal aspects of usability and progress indicators sought only to determine if users preferred progress indicators (Myers, 1985). The study showed that users did indeed prefer progress indicators, with the explanation that novice users view it with the understanding that the system is functioning normally for long tasks, while expert users use the progress indicator as a gauge allowing them to perform concurrent multiple tasks. From observation during the study, the participants who had the progress indicator watched the screen, whereas those who did not have a progress indicator looked around the room and stopped paying attention to the computer display. From a productivity perspective, the progress indicator is undoubtedly beneficial.

Generally, progress bars should be used for SRT more than ten seconds long (Nielson, 1994). Progress bars have three main advantages, (a) the user is aware that the system is still working, (b) the user can determine how much longer the wait will be, and (c) they provide visual interest which is directly correlated with preference (Myers, 1985). Visual interest is the reason why a graphic progress bar is preferable to displaying the remaining time in numbers, as well as the reason why graphical interfaces have won out over text-based user interfaces.

Recent research has been conducted on the way the progress bar behaves and which behavior users prefer (Harrison et al., 2007). Nine different progress bar behaviors were studied and users chose which behavior they preferred in a three-alternative study. The behaviors were each controlled by different non-linear functions and included power functions, wavy functions, and functions with pauses. The participants were presented with two of the nine possible progress bars and chose which one they preferred, or neither. The study found that participants perceived progress bars with pauses as taking longer to complete, and that an accelerating progress bar's behavior was strongly favored. However, all of the progress bars took 5.5 seconds to complete, which is not a long time to wait for a response from the computer, and research has shown that system SRT of that duration generally do not use a progress bar (Nielson, 1994). In addition, the progress bars were presented in an artificial way: the participants were simply asked to compare them. The results could have been different if the participants encountered the progress bars in a natural scenario, for instance while doing a task in which a SRT is present. In the experiment, the participants directly compared the behaviors of two different progress bars, which is seldom possible in normal computer use, and is not the goal of users viewing them. Few people compare progress bars they encounter and seek to decide which appears to take less time to complete. If the participants were working on another task while the SRT occurred, they would be responding to the effectiveness of the progress bar and not perceived duration of the SRT.

Motion and Time Prediction

Progress bars allow for estimating SRT duration the same way any moving object's arrival time at a future location is estimated, by prediction of the motion. Predicting the future location of a moving object requires three steps (Rosenbaum, 1975). First, a person must determine the rate of the motion and its direction. Next, a person must extrapolate that information to a specific time. Third, a response to the spatial and temporal motion must be started. A common example of this motion prediction is a person catching a ball. First, the catcher must determine the rate and trajectory of the ball, then, from that information, determine when it will reach him, and finally he must put himself into position to catch the ball. From this model, it has been suggested that a timing strategy is used to predict motion, with a slight modification on the three steps (Tresilian, 1995). However, research now shows that an attentional tracking strategy is used during motion extrapolation, and not a timing strategy. It has been shown that motion extrapolation is affected by spatial factors such as moving distractors (Lyon & Waag, 1995). Furthermore, people are able to extrapolate motion for objects that unpredictably disappear and reappear with changes in velocity (DeLucia & Liddell, 1998).

A similar series of events occurs when a person encounters a progress bar. First, the observer must determine the speed and behavior of the progress bar. Second, with that speed and acceleration information, the observer must determine at what time the progress bar will fill up, or when the SRT will end. Rosenbaum (1975) put forth a hypothesis of motion prediction of projecting the observed motion of an object via mental imagery. It was shown that observers perceive velocity and acceleration directly and accurately by mentally extending the motion. The participants were able to respond to acceleration and not just an average velocity. The findings suggested that the motion perception system is tuned to sense acceleration more readily than constant velocity. As the visual system is responsive to changes in stimulation (Caston & Bricout-Berthout, 1985; Cao, Gu, & Wang, 2004), and acceleration is the change of the change of position across time and constant velocity is the change of position across time, the findings are not surprising.

Additional research on acceleration perception found limitations to the visual system's ability. Werkhoven, Snippe, and Toet (1992) observed in a series of studies that the visual system is insensitive to acceleration over brief periods of time. Later experiments provided more evidence for the shortcomings of the visual system (Port, Lee, Dassonville, & Georgopoulos, 1997) and determined the minimum window to detect acceleration (Brouwer, Brenner, & Smeets, 2002). The minimum temporal window to detect a velocity change of 25% was determined to be 300 ms. This small window of time may be enough for observers to accurately assess a progress bar's behavior and know if it is accelerating, decelerating, or at a constant velocity. The current study sought to determine if there is a negative effect on performance in two-task circumstances because of acceleration and the minimum temporal window required to sense it.

Although the visual system is primed to recognize acceleration and changes in motion, our internal clocks behave linearly. Research on time estimation has shown that people are highly accurate at reproducing short time intervals—from 500 ms to 1,300 ms—and recognizing their ability by providing a judgment of their reproduction (Wearden and McShane, 1988).

People are also highly accurate at identifying longer durations—from 2 seconds to 8 seconds—even when engaged in a second task to prevent counting (Wearden, Denovan, Fakhri, and Haworth, 1997). These two studies indicate that people have a linear and accurate internal perception of time, even when faced with a task to hinder their performance. Further research on subjective versus real time confirmed that people do have a linear internal perception on time (Wearden and Jones, 2007) and no evidence was found to support a nonlinear representation of time.

In human-computer interaction, time can be represented in a variety of ways, each offering different benefits and shortcomings to the user. Therefore, time estimation is an important aspect of human-computer interaction. Two of the most prominent theoretical models of time estimation are the storage size model (Ornstein, 1969) and the change/segmentation model (Poynter, 1989). The storage size model dictates that time duration estimates are based on the amount of memory space that is required to store information about the time interval. Therefore, the more events or pieces of information occurring in the interval, the longer the duration will seem. The change/segmentation model states that the greater the degree of change that occurs in the time interval, the longer the duration will seem. From these two models, it can be inferred that if more discrete changes occur over a time period, then the duration of time that has passed will seem longer.

Time intervals are not always judged accurately and their estimation can be influenced by a number of nontemporal characteristics, including whether the time period is filled with a stimulus or not. Filled intervals are judged as lasting longer than unfilled intervals of the same duration (Craig, 1973; Steiner, 1968). In addition to time durations, this effect is present for visual stimuli. Known as the filled-space illusion, a line, area, or volume will appear larger if it is occupied by a number of distinct elements than if it is empty (Coren & Girgus, 1978). The filled-duration and space illusions are explainable by the storage size model of time estimation.

Time estimation can be assisted by using boundaries; the boundary is an indication of the start and stop of the time period. By giving users concrete beginnings and ends, the time interval can be more easily defined. A boundary can either be an explicit or implicit indication of the time passing. An example of an explicit boundary for a computer interface would be a bar filling up, and the end cap being the indication that the time interval is over. An example of an implicit boundary would be a percent-done indicator, which signifies to the person that when the percentage reaches 100, the task is completed. The boundary's existence communicates to the user that the task has an end and, if the task is completing uniformly, when that end will be reached. It has been shown that users prefer SRT representations in which a boundary is displayed over those without a boundary (Meyer, Bitan, & Shinar, 1995).

Purpose of the Research

Nonlinear graphical representations have had minimal research (Harrison et al., 2007), therefore the current experiment sought to expand on the previous limited research on nonlinear SRT representations. Additionally, the experiment sought to create an integration of the different representations and study how the representations affect users. The main goal of the current research was to determine which type of SRT representation is most beneficial in terms of task performance, specifically in allowing the user to return to working with the system most efficiently while accurately performing other concurrent tasks. Being able to engage multiple tasks efficiently is beneficial for many types of interfaces and situations. By integrating the findings of prior research studies, the best representation can be created and studied.

An additional goal of this study was to examine how the user's frustration level is affected by the different SRT representations. This is an important element in user interface design that affects user performance and is crucial to a user's experience with an interface. An increase in frustration with the interface can lead to poorer performance.

The current study also investigated whether a user can accurately assess the SRT duration by using the graphic representation of a determinate progress bar. A progress bar is a graphic representation of time, which is at the intersection of motion estimation theory and time estimation theory. The existing research on motion estimation would suggest that nonlinear progress bars would allow for optimum performance because of the innate ability of the visual system to detect acceleration. The research on time estimation would suggest that a linear progress bar would allow for optimum performance due to it's alignment with how our internal clocks behave, which is quite the opposite. The current research will provide insight on this unexplored intersection of these two theories.

The progress bars used in the experiment apply all of the recommended time affordances aside from exception, and in one case heartbeat. Acceptance is given by the progress bar appearing, and the affordance of scope is displayed graphically via the length of the progress bar and the rate of its fill, or its progress affordance. Initiation is indicated by the animation of the progress bar, and the affordance of remainder is displayed by the amount of the progress bar that is empty. The affordance of completion is achieved by displaying the result of the search. Exception is not needed because at no point during the SRT is user intervention required. A heartbeat is present for the continuous progress bar because it is constantly moving. However, a heartbeat is not present for the segmented progress bar because there are lengths of time where it appears that the progress bar is not doing anything. Without adding an additional cue or changing the segmented progress bar, this is a necessary component of the progress bar.

Due to the nature of progress bars and the many factors influencing them, this study focused on progress bars that were predictable and behaved with a single, constant of change (negative, zero, or positive). The experimental progress bars were carefully controlled and not influenced by external factors such as internet connection speed, other processes that were running, or unpredicted errors in the process. When all factors of the SRT are known and controlled, the SRT can be reliably measured.

Hypotheses. Based on the review of relevant literature, three hypotheses were formulated for further testing:

1. A continuous progress bar will have an advantage over a segmented one because it provides a higher resolution picture of the passage of time, allowing for quicker and more accurate estimation of the time remaining. These advantages may be quantified by shorter observation time of the progress bar before task switching, fewer intermediate checks of the progress bar, and more accurate return to the primary task. Based upon the literature on time estimation, a SRT representation that involves a continuous object should be seen as taking a shorter amount of time to complete, and therefore be preferred by users who will then be more efficient. This hypothesis is drawn from the storage size model of time estimation (Ornstein, 1969). Since time estimates are based on the amount of events or pieces of information occurring in the interval, the progress bar that is segmented into separate pieces should appear to take longer to complete and be judged less accurately.
2. Linear behavior of the progress bar will provide for best the performance due to the inherent ability people have for estimating time in a linear fashion (Wearden and Jones,

2007), contrary to the inherent ability of the visual system to perceive acceleration (incl. negatively accelerating, or decelerating) motion (Senot, Prévost, and McIntyre, 2003). Accelerating and decelerating progress bars will result in qualitatively different performance decrements, however; with an accelerating bar the participants will be late in returning to the delayed task, with decelerating bar too early, and in the latter case, resulting in additional task switching (Brouwer, Brenner, & Smeets, 2002).

3. Based on the previous progress bar behavior research, the bar that follows the power function should be preferred in a real world task. Harrison (2007) found that progress could be slowed in the beginning and accelerated towards the end, giving the user the illusion that the task is rapidly finishing, which was highly favored by users. Alternatively, the inverse power function should be least preferred because it will appear that the progress is slowing and the conclusion of the SRT is far off.

Experimental Tasks

The primary task required the participants to find and retrieve information from a database of words. The participant searched the database using a series of words given to them in the experiment packet. Upon entering a word, the program returned an associated word. Each of the words corresponded to a condition of the trial and each participant was exposed three times to each of the different conditions with a different word for each replicate. The order of conditions was randomized for each participant and one half of the participants saw the segmented progress bar section first, while the other half saw the continuous progress bar section first.

The secondary task was a visual search that required the participants to find a gray target square among black distracter squares. The target square did not appear in some of the searches and the participant was given the option to indicate that the target was not present. The secondary task trials were randomized independently from the first task.

Independent Variables

The independent variables used in this study were SRT duration, graphic representation of the SRT, and SRT behavior.

SRT duration. By varying the SRT, the different representations of the time delay could be compared across multiple time intervals within and between representations. In addition, encountering SRTs that were not uniform in duration was more characteristic of how systems actually behave. The SRTs used were 10 and 20 seconds. They represent common durations to wait for a system to respond, while adhering to Nielson's (1994) recommendation for the minimum duration of a SRT that should display a progress bar.

Graphic SRT representations. Two different SRT representations were used: a segmented determinate representation and a continuous determinate representation. The segmented determinate representation showed ten boxes filling up one by one until all were full and the SRT was over. The continuous determinate representation showed a single bar that filled up until the SRT was over (similar to Mac OS X determinate progress bar).

SRT Behavior. Seven different functions based on Harrison et al. (2007) were applied to how the progress bar behaved. The seven functions used were a linear function, three power functions, and three inverse power functions. The power functions used were $y = \left(x + \frac{(1-x)}{2}\right)^8$ (A3), $y = \left(x + \frac{(1-x)}{2}\right)^6$ (A2), and $y = \left(x + \frac{(1-x)}{2}\right)^2$ (A1), where y equals the displayed progress and x equals the actual progress, and the inverse power functions were $y = 1 + (1-x)^3 \times -1$ (D3),

$y = 1 + (1 - x)^{2.25} \times -1$ (D2), and $y = 1 + (1 - x)^{1.5} \times -1$ (D1). And the linear function was y = x (L).. The power functions represented three different accelerating and three different decelerating rates of change in the progress bar. Figure 1 is a graph comparing the seven different functions used.

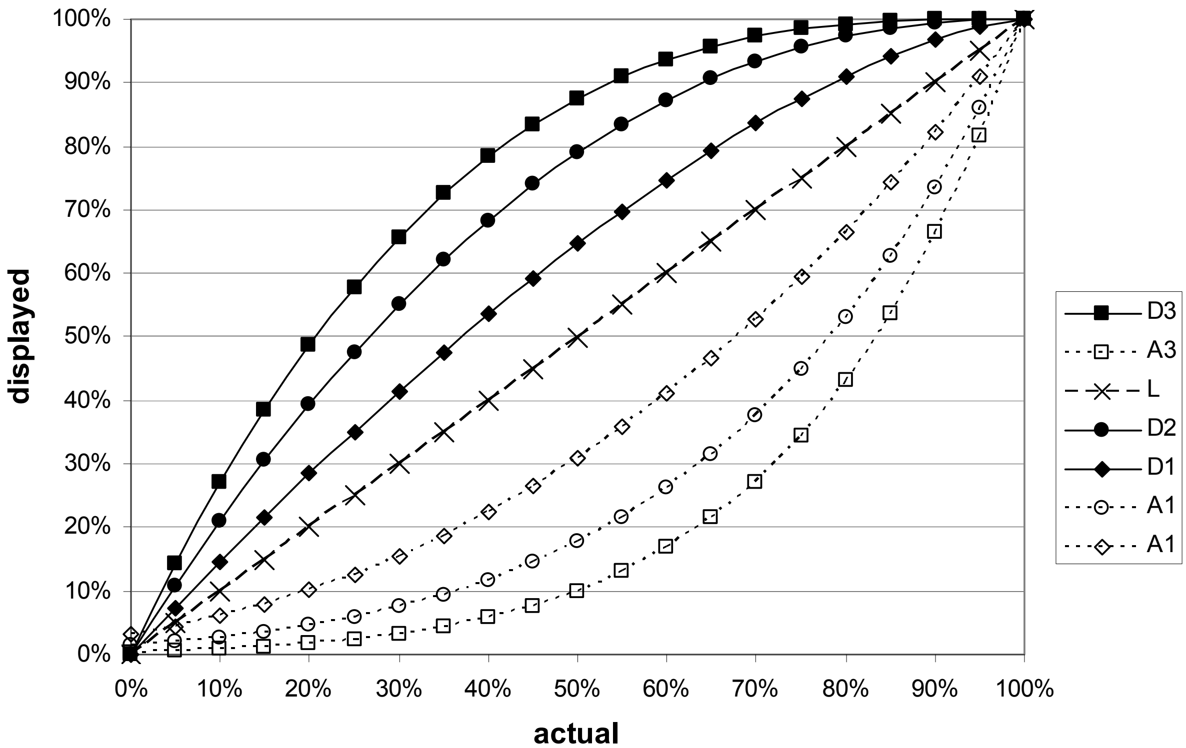


Figure 1. Comparison of the displayed and actual progress for each of the seven different progress bar behaviors used in the experiment.

Dependent Variables

The dependent variables measured were: (1) both a count of completed visual searches and speed of performance on the secondary task, (2) length of time before switching to the secondary task, (c) frequency of checking time left in the primary task as well as the length of each check, (3) timeliness of return to the primary task, measured as the time before or after the completion of the SRT, and (4) user satisfaction, frustration, and preferences of SRT representation, collected with the post-block and post-test questionnaires.

The post-block questionnaire used was a subjective questionnaire on user experience modified from the NASA Task Load Index (TLX) (Hart & Staveland, 1988). The NASA TLX is proven and dependable tool for measure subjective work load (Eggemeier, Wilson, Kramer, & Damos, 1991; Hendy, Hamilton, & Landry, 1993; Hitt, Kring, Daskarolis, Morris, & Mouloua, 1999), and was modified to meet the needs of the study. The questions used from the NASA TLX asked participants to rate how mentally demanding the task was, how hurried or rushed was the pace of the task, how successful they were in accomplishing what you were asked to do, how hard did they have to work to accomplish their level of performance, and how insecure, discouraged, irritated, stressed, and annoyed they were (Appendix A). The physical effort question was removed as no real physical effort was present in the study and the question would have been out of place. Also, the sources of load weighting were not used for the study, only the magnitude of load ratings was used. Additional questions asked the participants how well they thought the progress bar represented the time delay, using a similar seven point Likert scale, similar to the NASA TLX. The questionnaire also asked participants how long they felt the searches took to complete (in seconds). The post-block questionnaires were used to get subjective measurements of the two progress bar types. The post-experiment questionnaire asked the participants about which progress bar they preferred and which progress bar behavior they preferred. The post-test questionnaire also included questions on demographics, including the frequency of using a computer, experience installing programs or downloading files, sex, and age.

Chapter 2: Method

Participants

A total of 27 participants, 14 male and 13 female with a mean age of 20.5 years, volunteered for the experiment. All participants typically used computers at least once a day and all had experience downloading files or installing programs on their computers; hence, each participant was familiar with progress bars. No particular computer proficiency was required from the participants; the tasks were easy to complete for both novices and experts. Participants were recruited through friends, colleagues, and classes at RIT and they were motivated to perform their best with the incentive of a 50 dollar prize. Where applicable, students who participated also received extra credit in their psychology course.

Apparatus and Stimuli

The apparatus used in the experiment was an interface created specifically for the experiment. The interface itself was programmed in Adobe Flash and appeared to be the front-end interface for a search program. The interface only generated predefined responses to specific words. A search for any word that was not already defined in the program resulted in a "No results found" message. When the program was started, a timer was also started that kept track of how long the program had been running. When certain actions were performed by the user (search, switch tasks, complete a trial in task two) the timer was checked and the current time was recorded for that action. The computer used for the study was a Dell Optiplex GX260 with a 17-inch Dell Ultrasharp monitor running at a resolution of 1280x1024 pixels. Figure 2 shows two screenshots of the interface. The colors used in the study were all grayscale except for the progress bar. The progress bar was a gradient from a light green (H=61, S=147, B=194) to a dark green (H=61, S=138, B=113). For task two, the target was a dark gray (H=0, S=0, B=84), the distracters were black (H=160, S=0, B=0) and they were presented on a field of light gray (H=0, S=0, B=232). Appendix B contains the list of words used to search and the responses the system displayed.

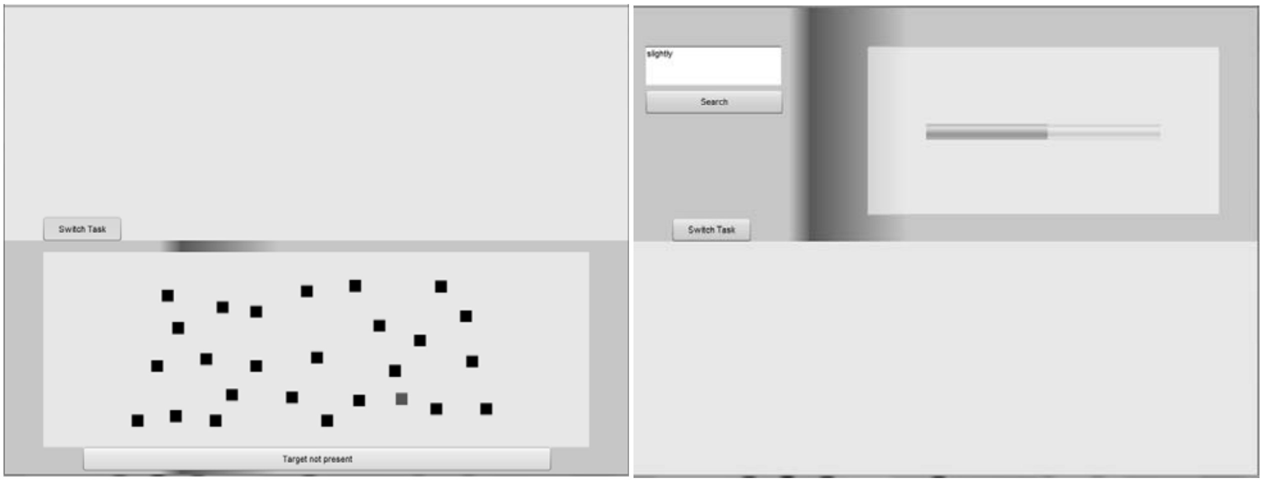


Figure 2. Screenshot of the interface, primary and secondary tasks. In the primary task, the participants were required to type a word in the search window and click the 'search' button. Upon the appearance of the progress bar, they were instructed to switch to the secondary task, which was a simple visual search task (in this example, the target is the grey square on the bottom row, third from the right) and complete as many trials of this task as they thought they had time for during the SRT of the primary task.

Procedure

The participants were given two tasks to complete in parallel, both being performed on the computer. The two tasks did not relate to one another, so that while the participants were waiting for the computer’s response on the first task, they could work on the second task without influencing the primary task.

Participants were motivated to do their best as they were instructed that whoever performs best would receive a $50 reward. They were told that they would be scored based on how well they do the two tasks. They were instructed that performing well on the primary task was crucial to their score. They were timed on how long it took to complete all of the searches for the primary task. For the secondary task, the total number of correct identifications of the gray box was recorded. The scores for both tasks were combined to give a final score that was used for the reward. The participants were informed that to obtain the best score, they should try for a low time on the primary task and a high number of correct identifications on the secondary task. The scoring function was $[0.2 \times (\text{ \% correct in task 2})] + [2000 - (\text{time for task 1})]$. This function weighted the time to complete the primary task higher, but made it worthwhile for the participant to strive to achieve many correct completions of the secondary task. The percentage of correct secondary tasks was used so that a participant could not continually click the "target not present" button to achieve a high score. The time to complete the primary task was subtracted from a standard score of 2000 so that a lower time on the primary task would result in a better score. Participants were not aware of the exact function used to weight the scores so that they could not pre-determine an optimal behavior or method to the task.

Once the first block of trials was completed (segmented or continuous progress bar), the participant was given a post-block questionnaire. The participant then completed the second block of trials and was given a second post-block questionnaire, followed by a post-experiment questionnaire (Appendix A). Once these questionnaires were completed, the participant was debriefed. Each block of trials required approximately 25 minutes to complete and the post-experiment questionnaire required 5 minutes, for a total experiment length of about 55 minutes.

Design

The design of the experiment was a 2 (progress bar type; continuous or segmented) x 2 (progress bar duration; 10 seconds or 20 seconds) x 7 (progress bar behavior; decelerating rate 1, decelerating rate 2, decelerating rate 3, linear, accelerating rate 1, accelerating rate 2, accelerating rate 3) within-subjects fully factorial design. In addition, each participant experienced each of the conditions three times for a total of 84 trials. The experiment was broken into two blocks, with one block containing the segmented progress bars and one block containing the continuous progress bars. The two types of progress bars were broken into blocks so that the participants could answer the two post-block questionnaires with minimal interruption. The data were compared across all conditions and participants to determine effects of progress bar type, duration, and behavior on user performance and preference.

Chapter 3: Results

For each condition the following metrics were recorded by the program: the length of time spent watching the progress bar before switching to task two; the number of checks on the progress bar (as well as the length of the check); whether the participant was early or late to return to the primary task with respect to the end of the SRT; or if they watched the entirety of the progress bar; the accuracy of their return to the primary task, with respect to the end of the SRT (mean error); the number of successful completions of the secondary tasks; and the responses to the questionnaires. The trials where the participant watched the entirety of the progress bar were not used for the analyses that use the metrics of first switch to the secondary task, number and length of checks, accuracy of return and performance on the secondary task. Because the participant never switched to task two in these trials, there are obviously no data available for these metrics. On average, participants watched six out of eighty-one trials.

During the administration of the test, a bug in the program was randomly present for some of the conditions causing an error in how the data were recorded for these conditions. This caused the program to record that they experienced fourth, fifth, and sixth replications of the condition and did not experience all the replicates for some of the other conditions. This error was present randomly across all conditions. The conditions that were recorded as the fourth, fifth, and sixth replicates were removed from all the analyses. Due to the nature of the problem, 26 trials per condition per replicate per participant (26 x 3 x 28) remained viable, for 2184 total usable data points and 84 discarded data points (3.8% lost data).

Preliminary Analyses

As a first analysis, each of the metrics was plotted as a histogram to determine the normalcy of the data and check for outliers (Appendix C). The counts of number of early returns and late returns, and number of times watching the entirety of the progress bar was also plotted by condition as well as replicate.

Data transformations. Visual inspection of the histograms of the dependent variables indicated that length of time before switching to the secondary task was highly positively skewed, as is typical to time data (Figure C1). To restore normality of the distributions, a base ten log transformation was done to these data (Figure C6).

Table 1 summarizes the dependent variables measured for each of the 28 conditions, and table 2 summarizes the results by independent variable for ease of comparison. Initially, 2 (SRT duration; 10 second and 20 second) x 2 (Progress bar type; continuous and segmented) x 7 (Progress bar behavior; decelerating rate 1, decelerating rate 2, decelerating rate 3, linear, accelerating rate 1, accelerating rate 2, and accelerating rate 3) repeated measures ANOVAs were performed to examine the effects of progress bar type, SRT duration, and bar behavior on each of the dependent variables. Tukey's post-hoc comparisons were used to examine the differences between each level of progress bar behavior. There were no differences between the three accelerating behaviors or the three decelerating behaviors in any of the analyses. Hence, to observe differences between the three rates of change (accelerating, linear, or decelerating) the three accelerating conditions (A_1, A_2, and A_3) were averaged into one accelerating condition and the three decelerating conditions (D_1, D_2, D_3) were averaged into one decelerating condition. Also by averaging the accelerating and the decelerating conditions, the difference in the number of observations for each rate of change was made even. These remaining three conditions represented an accelerating, linear, and decelerating behavior, respectively. Subsequently, only 2 (SRT duration; 10 second and 20 second) x 2 (Progress bar type; continuous and segmented) x 3 (Progress bar rate of change; accelerating, linear, and decelerating) repeated measures ANOVAs were performed for all dependent variables. Although not used for analysis, appendix D includes all 2 x 2 x 7 ANOVA tables for reference. Appendix E contains all the 2 x 2 x3 ANOVA tables and appendix F contains all the post-hoc tables for both the 2 x 2 x7 and 2 x 2 x 3 ANOVAs.

Outliers. For length of time before switching to task two, four data points were determined to be outliers as the length of time before the participant switched to task two was nearly identical to the length of the SRT, indicating that these participants did not do the tasks as instructed. These four data points were over four standard deviations away from the condition's mean and were removed from the dataset. The other dependent variables did not contain any outliers that would potentially distort the data. Appendix I contains boxplots of the averaged data.

Table 1

The complete results of the experiment.

Condition	Time Until First Switch to Task Two		Count of Number of Checks			Length of Check		Accuracy of Return to Task One		Count of Early and Late Returns, and Watched SRTs			Number of Completions on Task Two	
	Mean	SD	0 checks	1 check	≥2 checks	Mean	SD	Mean	SD	Late	Early	Watch	Mean	SD
10 Second SRT														
D3 CON	1391.03	1557.81	38	29	11	836.13	430.13	2141.96	6572.54	38	32	8	4.59	4.11
D3 SEG	1799.40	2210.32	48	21	9	809.41	536.29	1480.55	3109.46	43	24	11	4.21	3.05
D2 CON	1405.45	1567.10	46	28	4	754.35	348.32	373.37	3898.39	26	41	11	4.10	2.52
D2 SEG	1747.35	2055.22	49	22	7	792.53	412.70	1206.77	3999.61	45	24	9	4.62	3.25
D1 CON	1506.33	1882.82	59	16	3	725.08	189.47	2167.29	4249.42	49	25	4	5.29	3.18
D1 SEG	1825.63	2128.59	53	20	5	744.93	353.18	812.25	3943.31	39	29	10	4.21	2.93
L CON	1186.22	1512.11	67	11	0	487.18	139.09	5010.66	7401.40	55	19	4	6.96	4.47
L SEG	1626.24	2002.22	70	8	0	582.38	239.89	3749.42	4185.37	66	6	6	6.33	3.85
A1 CON	1378.51	1675.41	66	12	0	587.92	265.50	5556.29	6124.77	58	11	9	6.77	4.50
A1 SEG	1450.49	1607.56	64	13	1	648.36	342.15	5689.05	6224.07	65	8	5	7.22	4.52
A2 CON	1235.13	1452.39	68	10	0	901.40	1187.09	5707.22	6692.93	57	15	6	7.19	4.45
A2 SEG	1619.29	1968.56	68	10	0	487.90	67.86	5270.99	5489.04	63	8	7	6.27	3.92
A3 CON	1322.59	1466.53	68	10	0	445.20	72.80	7829.38	8146.91	63	8	7	7.53	4.89
A3 SEG	1587.93	2016.33	61	15	2	517.18	93.71	5284.80	4602.80	70	4	4	7.32	3.26
20 Second SRT														
D3 CON	1860.92	2775.58	15	31	32	964.77	513.68	1198.84	4246.51	29	46	3	7.62	3.77
D3 SEG	2168.13	2951.64	25	23	30	1084.89	764.29	788.00	4531.25	34	41	3	8.09	4.46
D2 CON	1985.74	2592.21	27	20	31	1087.98	648.14	230.10	5134.52	28	42	8	7.18	4.89
D2 SEG	1983.65	2341.96	23	24	31	1154.85	848.93	42.08	4777.78	35	36	7	7.19	4.11
D1 CON	1576.80	2146.17	29	22	27	817.41	327.46	314.51	3800.43	33	38	7	8.55	4.70
D1 SEG	1843.20	2629.35	20	31	27	808.83	367.55	906.80	4718.08	44	30	4	9.04	4.14
L CON	1347.36	1648.62	29	32	17	852.99	670.30	1838.16	4461.02	45	28	5	9.69	4.34
L SEG	1859.76	2561.62	22	32	24	785.84	428.62	1980.91	3944.05	51	23	4	9.36	4.67
A1 CON	1381.21	1707.06	25	42	11	743.95	345.37	3715.81	8796.23	53	20	5	10.88	5.69
A1 SEG	2081.34	3367.00	23	35	20	862.79	964.63	3431.88	6150.17	61	16	1	11.13	4.61
A2 CON	1853.60	2819.85	28	35	15	761.07	344.34	3865.35	5715.92	63	12	3	11.08	5.81
A2 SEG	1899.24	2494.79	24	38	16	872.50	436.06	3977.55	4587.91	65	11	2	11.27	4.24
A3 CON	1462.70	1968.70	28	32	18	743.10	345.19	5287.72	6916.57	61	15	2	11.62	5.01
A3 SEG	2035.14	2820.72	21	41	16	984.02	1376.19	4621.66	6036.94	67	9	2	11.26	5.25

Note. Times listed in ms

Table 2
The complete results of the experiment, comparison by independent variable with progress bar behaviors averaged.

Progress Bar Behavior	Progress Bar Type	Time Until First Switch to Task Two		Average Number of Checks		Length of Check on Primary Task		Number of Completions on Task Two		Count of Early and Late Returns, and Watched SRTs			Error of Return	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Early	Late	Watch	Mean	SD
10 Second SRT Duration														
Decelerating	Continuous	1559.07	1669.13	0.47	0.43	798.14	311.77	4.66	2.54	98	113	23	3151.45	3165.68
Decelerating	Segmented	1859.06	2003.34	0.47	0.50	825.47	474.38	4.34	2.38	77	127	30	2821.75	2147.93
Linear	Continuous	1186.22	1512.11	0.14	0.35	487.18	139.09	6.96	4.47	19	55	4	5561.20	6991.47
Linear	Segmented	1626.24	2002.22	0.10	0.31	582.38	239.89	6.33	3.85	6	66	6	4071.92	3867.75
Accelerating	Continuous	1360.73	1503.67	0.14	0.25	630.06	404.45	7.16	3.77	34	178	22	6779.10	5687.26
Accelerating	Segmented	1640.11	1790.38	0.19	0.32	588.27	267.74	6.93	3.14	20	198	16	5662.79	4349.43
20 Second SRT Duration														
Decelerating	Continuous	1859.94	2089.27	1.25	0.86	995.49	484.76	7.78	3.64	126	90	18	3247.44	2362.32
Decelerating	Segmented	1993.43	2199.44	1.30	1.00	991.58	451.07	8.13	3.26	107	113	14	3490.50	2228.59
Linear	Continuous	1347.36	1648.62	0.86	0.78	852.99	670.30	9.69	4.34	28	45	5	2947.07	3810.60
Linear	Segmented	1859.76	2561.62	1.10	0.91	785.84	428.62	9.36	4.67	23	51	4	3211.31	3013.21
Accelerating	Continuous	1612.92	1949.86	0.86	0.62	770.22	285.54	11.19	4.35	47	177	10	4954.21	4929.72
Accelerating	Segmented	1980.69	2525.32	1.02	0.75	1065.95	1406.40	11.22	3.35	36	193	5	4915.06	4833.35

Note. Times listed in ms

First Switch to Secondary Task

A within-subjects repeated measures ANOVA was performed to test the possible nuisance effects of replicate. Mauchly's test indicated that the assumption of sphericity had not been violated, χ^2 (2) = 1.65, p > .05. The results show that there were no significant differences between replicates, F(2, 13.73) = 1.86, p > .05 (Fig 3).

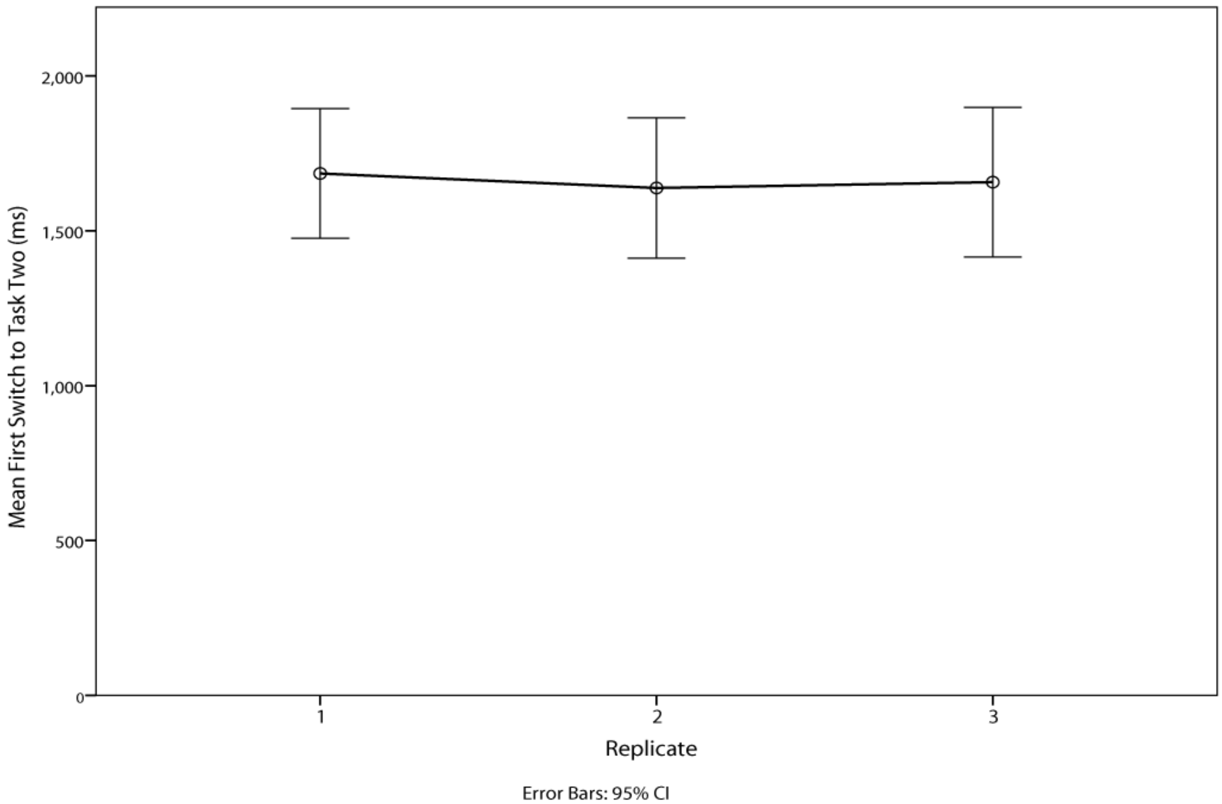


Figure 3. Effect of replicate on amount of time until first switch to the secondary task in ms. The graph indicates a possible trend of taking longer to switch to the secondary task for the first replicate, but this is not significant.

The data were then analyzed with a 2 x 2 x 3 repeated measures ANOVA (Table E1). There was a significant effect of progress bar type, F(1, 244) = 13.54, p < .05, with the continuous progress bar resulting in the shortest length of time before switching to task two. The mean difference in times was 338.89 ms. There was also a significant effect due to participants F(26, 244) = 51.36, p < .05, indicating a difference between each of the participant's results. Bar behavior was significant as well, $F(2, 244) = 8.21$, $p < .05$. Tukey *post-hoc* comparisons were used to determine differences between the progress bar behaviors (Table F6). The linear condition (M_L = 1504.60 ms) resulted in a significantly shorter amount of time before switching to the secondary task than the decelerating condition (M_D = 1818.32 ms) and accelerating condition (M_A = 1651.73 ms), $p < .05$. The difference between the linear and accelerating behaviors was not significant, $p < .05$. The decelerating behavior was not different from the accelerating behavior, $p < .05$. No difference was found between SRT durations, $p > .05$. No interactions between the independent variables were significant, either with all $p > .10$. Figure 4 depicts the differences between the conditions.

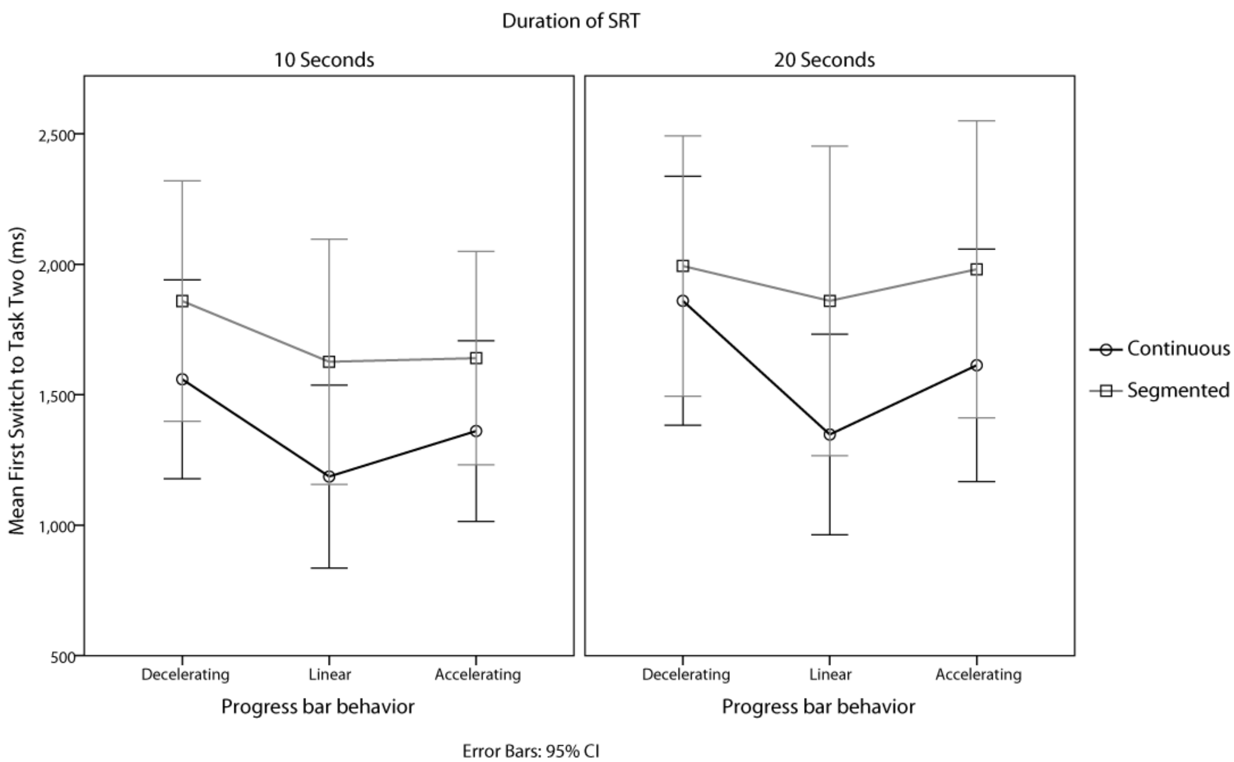


Figure 4. Effects of progress bar types and behaviors and SRT duration on the time before switching to the secondary task. The continuous bar and linear behavior allowed for significantly faster switches. No effect was found for duration of the SRT and no interactions were present.

Checks on Primary Task

Number of checks. Effects of the progress bar types and behaviors as well as SRT durations on the number of intermediate checks on the bar during secondary task performance were analyzed with a 2 x 2 x 3 repeated measures ANOVA for both the number of checks and length of the checks on the primary task (Table E2). Mauchly's test indicated that the assumption of sphericity had been violated, χ^2 (2) = 12.45, $p < .05$, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .96$). The results show that there were significant differences between replicates, $F(1.94, 532.34) = 6.079$, $p < .05$, indicating differences in the number of checks on the progress bar within the replicates. The mean values indicate that participants checked more often on the first replicate than the second and third (M_{R1} = 0.68, M_{R2} and M_{R3} = 0.65). There was a significant effect of bar type, $F(1, 274) = 8.07$, $p < .05$; the segmented progress bar resulted in the most number of checks ($M_S = 0.70$, $M_C = 0.62$). There was also a significant effect due to participants $F(26, 274) = 15.82$, $p < .05$, indicating a difference between each of the participant's results. Bar behavior was significant, too, $F(2, 274)$ = 27066, $p < .05$. Tukey post-hoc comparisons (Table F7) indicated that the accelerating and linear condition ($M_L = 0.55$, $M_A = 0.55$, respectively) resulted in a significantly fewer checks than the decelerating condition ($M_D = 0.87$), with $p < .05$. The SRT of 20 seconds also resulted in more checks ($M_{20} = 1.07$) than the SRT of 10 seconds ($M_{10} = 0.25$), $F(1, 274) = 396.88$, $p < .05$, which was not surprising. No interactions between the independent variables were significant, with all $p > .25$. Figure 5 depicts the differences between the conditions.

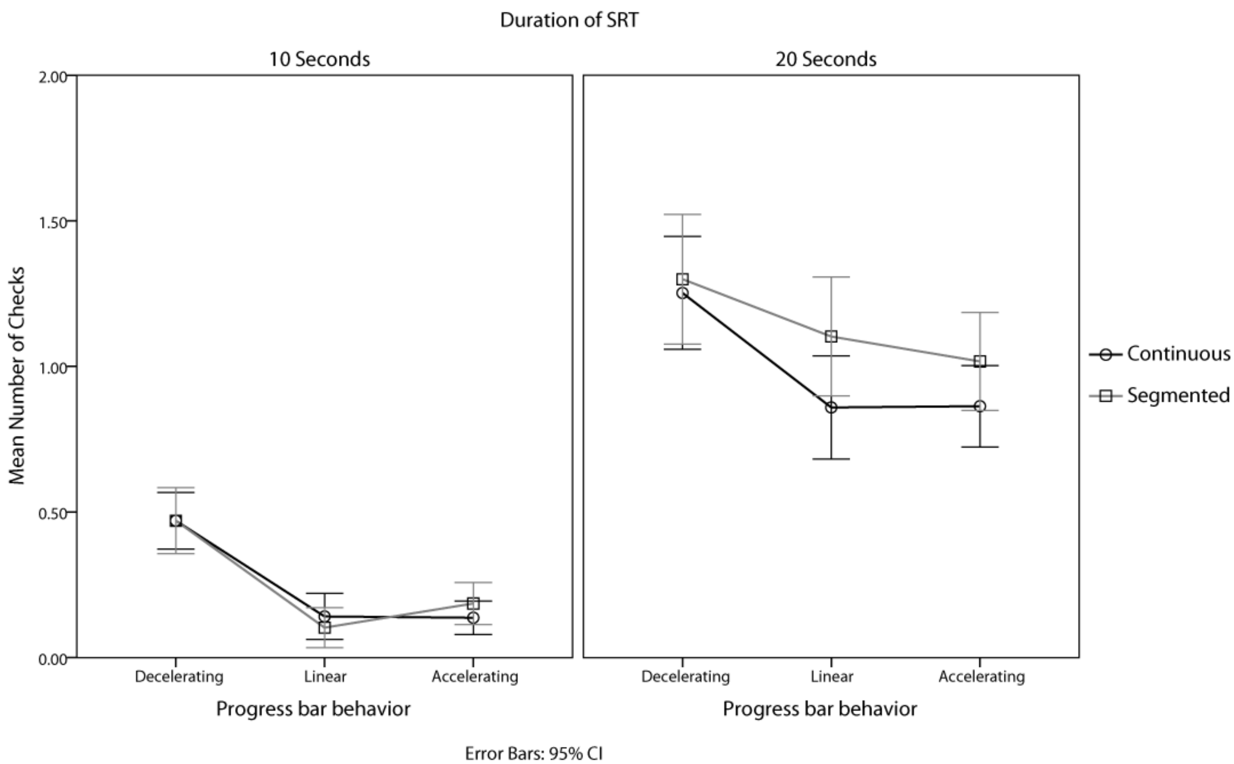


Figure 5. Effects of progress bar types and behaviors and SRT duration on the number of checks on the bar per trial. The 10 second SRT durations had significantly fewer checks and the decelerating progress bar behavior had significantly more. Duration of the SRT had a significant effect on the mean number of checks as well and no interactions were present.

Length of checks. The effects of progress bar type and behavior and SRT duration on the length of check on the bar before switching back to task two were analyzed similarly (Table E3). Mauchly's test indicated that the assumption of sphericity had been violated, χ^2 (2) = 90.72, $p <$ .05, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (ε = .62). The results show that there were no significant differences between replicates, F(1.23, 115.84) = 3.85, $p < .05$. There was a significant effect due to participants F(25, 94) = 4.28, $p < .05$, indicating a difference between each of the participant's results. There was not a significant effect of progress bar behavior F(2, 94) = 1.14, $p > .05$. Tukey post-hoc comparisons were used to determine differences between the progress bar behaviors (Table F8).

The comparisons of the three behaviors indicated that the accelerating and the decelerating conditions (M_A = 844.46 ms, M_D = 914.67 ms) did not differ, but resulted in significantly longer checks than the linear condition (M_L = 772.76 ms), with $p < .05$. The progress bar type, $F(1, 122) = 0.18$, $p > .05$, and SRT duration, $F(1, 122) = 0.75$, $p > .05$, had no significant effects on check durations. A significant interaction effect was present for the independent variables of SRT duration and progress bar type, $F(1, 94) = 5.90$, $p < 0.5$. See Figure 6 for the progress bar effects on length of checks.

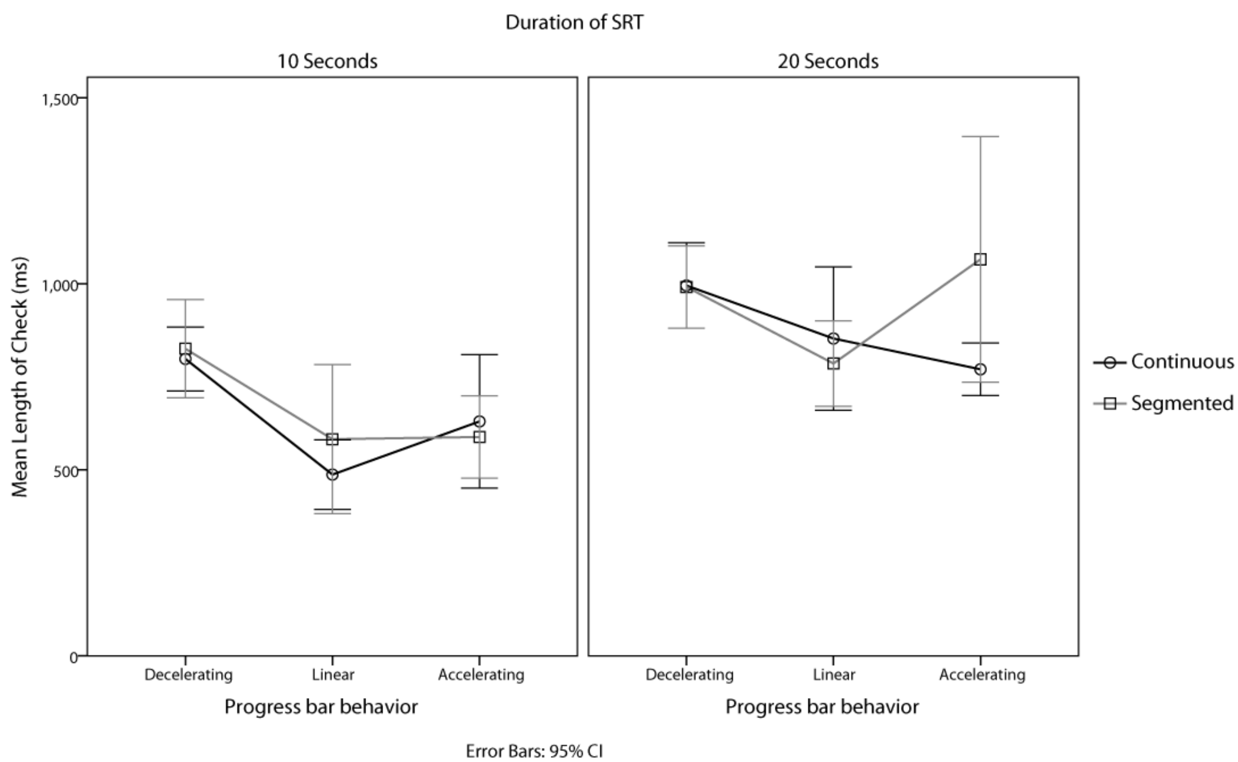


Figure 6. Effects of progress bar types and behaviors and SRT duration on the length of checks on the bar before switching back to the secondary task. There were no significant effects for each of the main variables, but a significant interaction was present for SRT duration and progress bar type indicating that the longer the duration and the segmented bar resulted in longer checks.

Accuracy of Return to Primary Task

Accuracy. A 2 x 2 x 3 repeated measures ANOVA (Table E4) was performed to test the effects of the independent variables on the accuracy of return to the primary task, measured by the time the participants were too early or too late returning after the SRT was complete (Error of return). Mauchly's test indicated that the assumption of sphericity had been violated, χ^2 (2) = 7.69, $p < .05$, therefore degrees of freedom were corrected using the Greenhouse-Geisser estimate of sphericity ($\varepsilon = .97$). The results show that there were no significant differences between replicates, $F(1.94, 473.25) = 1.39$, $p > .05$. There was a significant participant effect $F(26, 244) = 11.85$, $p < .05$, indicating a difference between each of the participant's results. There was not a significant effect for bar type, $F(1, 244) = 2.11$, $p > .05$. There was a significant effect of behavior, $F(2, 244) = 23.87$, $p < .05$ and duration, $F(1, 244) = 11.54$, $p < .05$. Tukey post-hoc comparisons (Table F9) indicate that the decelerating and linear conditions (M_D = 3179.99 ms, M_L = 3950.44 ms), $p < .05$, resulted in significantly more accurate returns than the accelerating condition (M_A = 5569.51 ms), but did not significantly differ from each other. The 20 second SRT (M_{20} = 3806.7 ms) resulted in more accurate returns to the primary task than the 10 second SRT did (M_{10} = 4675.56 ms). A significant interaction between the independent variables of SRT duration and progress bar behavior, $F(2, 244) = 6.25$, $p < .05$. See Figure 7 for the progress bar effects on accuracy of return to the primary task.

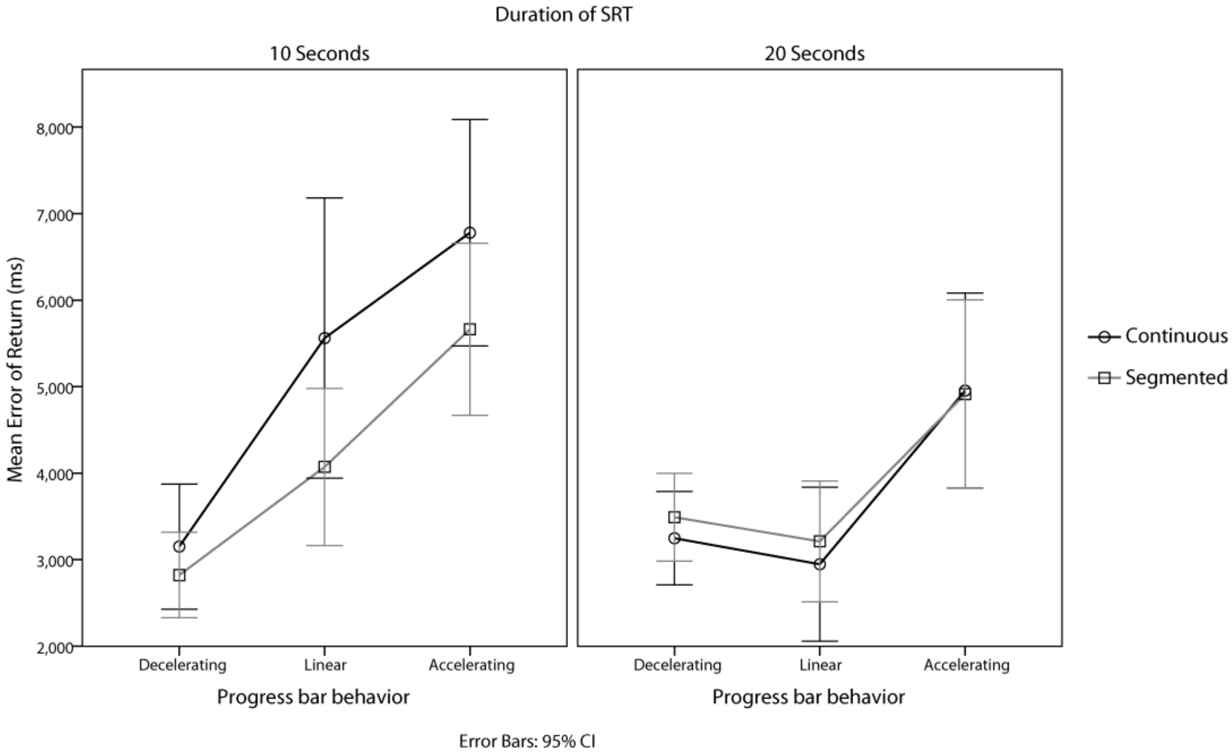


Figure 7. Effects of progress bar types and behaviors and SRT duration on error of returning to the primary task at the end of the SRT. Both progress bar behavior and SRT duration had a significant effect on the error of returning to the primary task, with the accelerating progress bar and shorter SRT duration resulting in greater error in returning to the primary task. An interaction was present for these two variables as well, as shown by the greater error in returning to the primary task for the 20 second accelerating progress bar.

To determine if a relationship between the number of checks (M = .66) and accuracy of return (M = 4238.24 ms) existed, these two metrics were first plotted against each other in a binned scatter plot (Figure 8). The bin size for they-axis is 1333.33 ms and the bin size for the x-axis is .32. A Pearson correlation was then performed and a significant relationship was found, $r(900) = -0.32$, $p < .05$. These results indicate that a relationship between progress bar behavior and number of checks was present, indicating that as participants checked more often they were more likely to return accurately to the primary task.

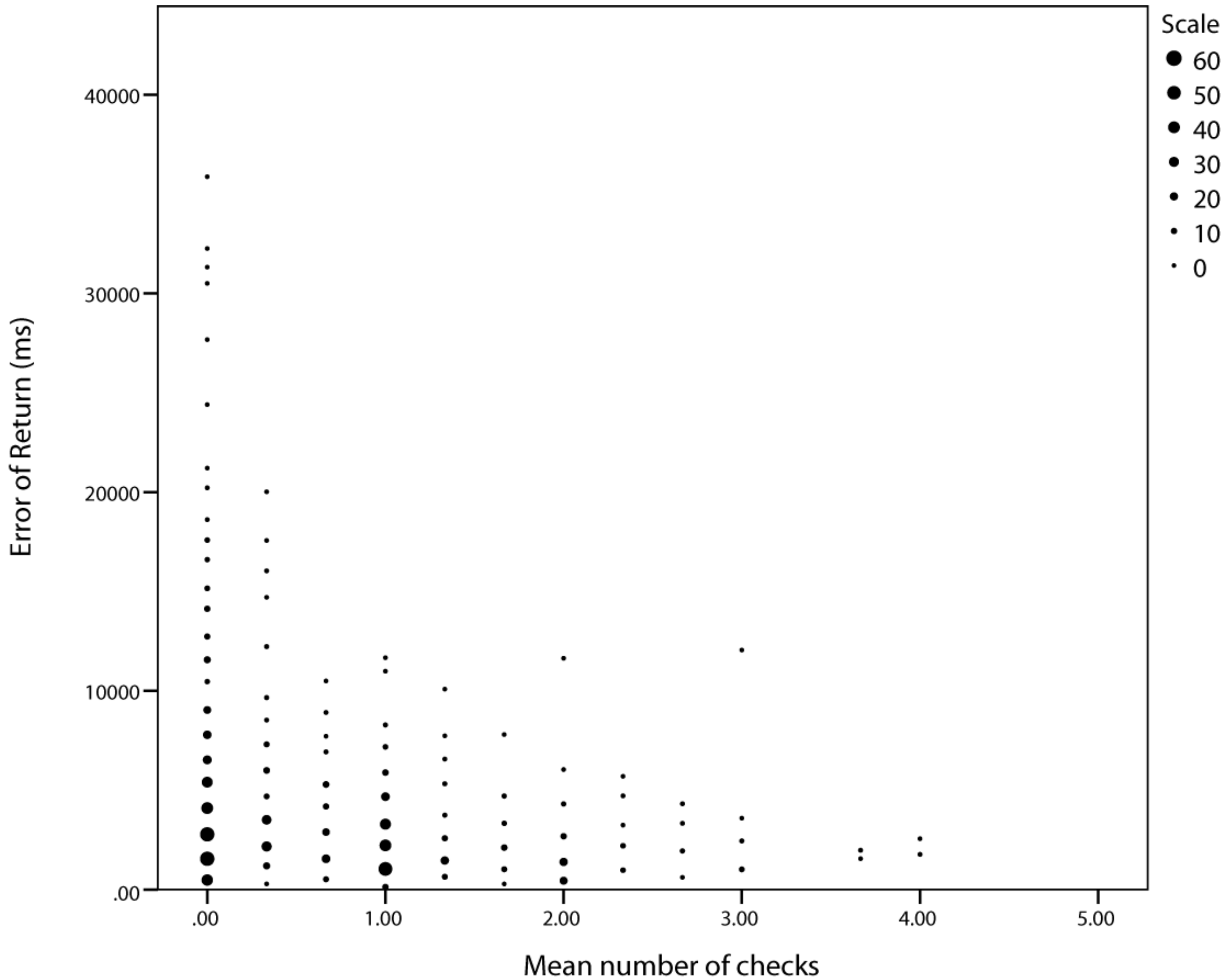


Figure 8. Comparison of number of checks on progress bar with error in returning to primary task at the end of the SRT. The size of the point represents the number of observations for that bin. The graph suggestions a relationship between the error of return and the number of checks on the progress bar, with more error of return related to fewer checks on the progress bar.

Proportions of early and late returns. To look for a pattern in the number of early and late returns to the primary task, the percentages for each were plotted against the three progress bar behaviors (Figure 9). The graph implies a strong trend of late arrivals as the behavior of the progress bar increases in the positive (accelerating) direction.

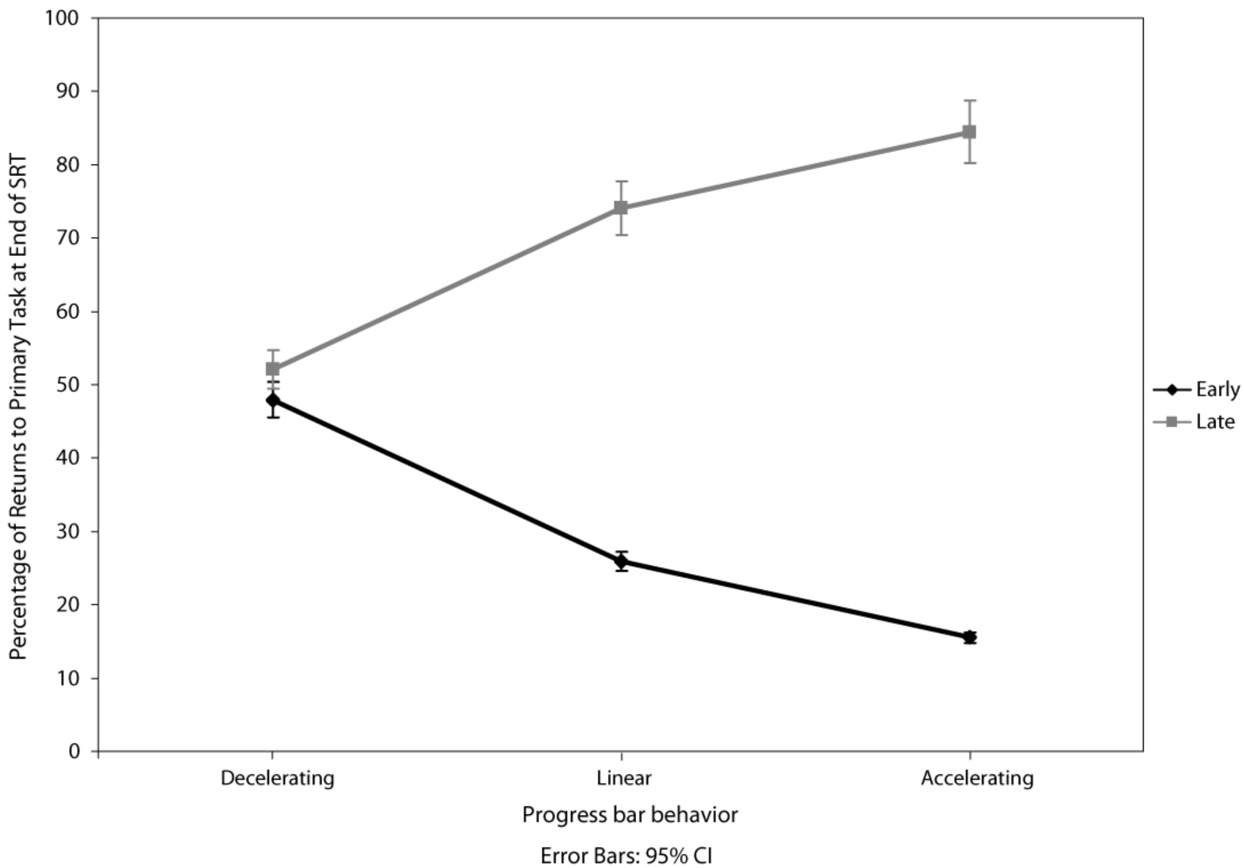


Figure 9. Percent of early or late return to task one with respect to progress bar behavior. The graph suggests a trend of an increasing percentage of late returns to the primary task as the progress bar acceleration increases.

Secondary Task Performance

To examine the effect of the experimental conditions on performance in the secondary task, measured by number of successful completions of the visual searches, a 2 x 2 x 3 repeated measures ANOVA was performed (Table E5). Mauchly's test indicated that the assumption of sphericity had been violated, χ^2 (2) = 8.15, $p < .05$, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .97$). The results show that there were significant differences between replicates, $F(1.94, 532.34) = 6.08$, $p < .05$. There was a significant effect due to participants $F(26, 274) = 59.44$, $p < .05$, indicating a difference between each of the participant's results. There was a significant effect of progress bar behavior, F(2,

274) = 49.24, $p < .05$. Tukey post-hoc comparisons (Table F10) indicated that the decelerating condition ($M_D = 6.23$) resulted in significantly fewer completions of task two than the linear condition ($M_L = 8.09$), which resulted in significantly fewer completions than the accelerating condition ($M_A = 9.12$), with $p < .05$. SRT duration was significant, unsurpsingly, $F(1, 274) = 209.39$, $p < .05$; clearly, 20 s ($M_{20} = 9.56$) SRT allowed for far more searches to be completed than the 10 s SRT ($M_{10} = 6.07$). Progress bar type was not significant (1, 274) = 1.29, $p > .05$, nor were there any significant interactions, with all $p > .05$.

Preferences and Subjective Ratings

A Chi-Square test for equal proportions was performed on the preference results for continuous or segmented progress bars. The progress bar type included 17 participants who preferred the continuous type and 10 participants who preferred the segmented type. These proportions were not significantly different, χ^2 (1, $N = 27$) = 1.815, $p > .05$.

Preferences for decelerating, linear, or accelerating progress bars were analyzed similarly. Two participants preferred the decelerating behavior, 18 preferred the linear behavior, and 6 participants preferred the accelerating behavior. These proportions were significantly different, with χ^2 (1, $N = 26$) = 16.00, $p < .05$.

The preference ratings were each summed for progress bar type and behavior. The progress bar preference included 0 participants who preferred the Segmented Decelerating progress bar, 7 participants who preferred the Segmented Linear behavior, 3 participants who preferred Segmented Accelerating, 2 participants who preferred the Continuous Decelerating, 11 who preferred the Continuous Linear, and 3 who preferred the Continuous Accelerating. These frequencies were significantly different, χ^2 (5, $N = 26$) = 18.33, $p < .05$. From these results, it is apparent that the progress bar type that participants preferred most was the continuous linear.

Questionnaires. To determine if significant differences existed between the questions on the continuous block questionnaire and the segmented block questionnaire a paired samples t-test was performed. Only the responses on how mentally demanding was the task was significantly different, $t(27) = 2.39$, $p > .05$, with the segmented progress bar (M = 2.93) being rated as less mentally demanding than the continuous progress bar (M = 3.43). The results of the post-block questionnaires and *t*-tests are presented in table 4.

Table 3
Post-block questionnaire results

Question	Type	Mean	N	SD	Std. Error Mean	*t*	df	Sig. (2-tailed)
How mentally demanding was the task? 1 = very low	CON	3.43	28	1.501	.284	2.393	27	0.02*
	SEG	2.93	28	1.464	.277			
How hurried or rushed was the pace of the task? 1 = very low	CON	4.39	28	1.729	.327	-.583	27	0.56
	SEG	4.54	28	1.598	.302			
How successful were you in accomplishing what you were asked to do? 1 = perfect	CON	2.93	28	1.562	.295	-.120	27	0.91
	SEG	2.96	28	1.688	.319			
How hard did you have to work to accomplish your level of performance? 1 = very low	CON	4.00	28	1.563	.295	.570	27	0.57
	SEG	3.89	28	1.663	.314			
How insecure, discouraged, irritated, stressed, and annoyed were you? 1 = very low	CON	2.89	28	1.499	.283	1.000	27	0.33
	SEG	2.68	28	1.541	.291			
I thought the progress indicator did a satisfactory job of representing the time delay 1 = completely agree	CON	3.36	28	2.004	.379	.795	27	0.43
	SEG	3.07	28	1.654	.313			
How long did you feel the searches took to complete (in seconds)?	CON	11.88	28	5.797	1.096	.474	27	0.64
	SEG	11.54	28	6.462	1.221			

* significant results, $p < 0.05$

Chapter 4: Discussion

In general, participants were affected by progress bar behavior and type across all metrics. With short observations on the progress bar behavior, participants are able to sense positive or negative acceleration and adjust behavior based on their observation. The length of time before switching to task two and the length of checks on the progress bar were all well above 300ms—at least twice that amount—therefore the minimum temporal window proposed by Brouwer et al. (2002) was maintained.

In some cases, participants chose to watch the entirety of the progress bar. This was not biased towards the first replicate or any specific condition. Therefore, participants did not use the first replicate of each condition to get a sense of how long the SRT would be. Likewise, they did not watch the final replicates to try to finish the experiment as soon as possible, which might have indicated fatigue in the participant, if it had occurred. The random distribution of these watched progress bars suggests that participants were occasionally watching the progress bar to check how long it was taking, or to take a break from the second task.

For most measures there was an effect due to participants indicating a significant difference between the participants. This is an expected result of the study, as it is assumed that each of the participants used their own strategy of performing the two tasks and the likelihood of all of the strategies being the same is improbable. The purpose of the study was not to look at how individuals performed differently, but to determine how the progress bars affected performance in general.

This study looked at progress bars that behave in a constant, predictable motion, which is seldom how actual progress bars behave in a system that is influenced by external factors. External factors are what make accurate SRT prediction difficult; other running processes, connection speed, and complexity of the task are just three things that make SRT predictions unreliable. In closed systems or where the SRT can be reliably predicted, the recommendations of this study can be applied. Again, predictability being a crucial aspect of SRT, as discovered in the prior research (Shneiderman, 1987). This study further confirms the necessity of predictability as well as feedback during the delay. Throughout all SRT in the study there was adequate feedback (Rushinek & Rushinek, 1986). It is most likely because of these two factors, predictability and feedback, the participants experienced little stress as described in the next section.

Task Performance and Preference

Bar Type. The first hypothesis that a continuous progress bar will allow for a more accurate SRT representation leading to better performance and higher preference was supported. The continuous progress bar allowed for better performance on the metrics of length of time before switching to task two, number of checks, length of check, and length of time between the end of the SRT and returning to task one; however, only the amount of time before switching to task two and number of checks was significantly better. Although the continuous progress bar did result in better performance for length of checks and accuracy of return to the primary task, these results were not significant

One possibility of why the length of time before switching to task two was shorter for continuous progress bars is that a segmented progress bar forces the user to watch until the first segment is in place. A continuous progress bar allows the user to get a quick and rough estimate of how fast the progress bar is filling by providing a constantly updating representation of the SRT. The segmented progress bar, on the other hand, provides no information to the user until the time required to fill the first segment has passed. The user must witness at least two segments to gain any knowledge on the rate of change of the progress bar. The number of segments present in the bar controls how quickly the user can receive the information necessary to make an initial estimate of how long the SRT will be. In this study, ten segments were present which, in the linear condition, forced the user to wait until one tenth of the SRT had passed. In conditions where the progress bars followed the power function (accelerating behavior), that initial time period was increased even further to a maximum of 3.2 seconds in the ten second SRT, which is almost one-third the total time.

Only for the initial judgment on the progress bar, and not on any checks on the progress bar, did participants watch long enough to gain a sense of the rate of change. No difference was found between the progress bar types for length of check, which indicates that the rate of change information was not accessed during checks. It appeared participants were only interested in checking to see if the SRT had ended. If the rate of change information was desired, the participants would have to watch the progress bar longer for the segmented bar than the continuous bar which would result in a difference between the two progress bar types. This indicates that participants are not checking and modifying their estimation on the motion (Brouwer, Brenner, & Smeets, 2002; Cao, Gu, & Wang, 2004; Caston & Bricout-Berthout, 1985) of the progress bar, but simply checking their initial time estimation (Wearden and Jones, 2007). It may be that people generally have greater trust in their estimations of motion than their estimations of time, resulting in the need to check the time estimation, but not the motion estimation.

Although there was a difference between the perceived durations of the continuous and segmented progress bars (M_C = 11.88 s, M_S = 11.54 s), the difference was not significant. The results do not support the storage size model (Ornstein, 1969) and the change/segmentation model (Poytner, 1989) of time estimation, nor do they invalidate it. Therefore, it cannot be determined if the perceived duration of the SRT was affected by progress bar. The results do contradict the theories on filled intervals (Craig, 1973; Steiner, 1968), as the progress bar that contained more discrete units was not seen as lasting longer than the progress bar with one discrete unit. In fact, they were judged to be almost identical.

Participants also favored the continuous progress bar over the segmented bar, which was also the progress bar that resulted in better performance. This supports the theory that users will prefer those displays that they perform well with and therefore do not cause frustration. Although not directly questioned, it is possible that the missing heartbeat of the segmented progress bar caused users to be less pleased with the segmented SRT representation. Additionally, participants did not express any difference in temporal demand, performance, effort, and frustration. But a small difference in mental demand was expressed, with the segmented bar being slightly less demanding, even though the participants did prefer the continuous bar. The apparent subjective similarity and objective dissimilarity between the two progress bar types indicates that although the participants are performing better with the continuous progress bar, there are no perceived differences to the participants.

As mentioned earlier, it is likely that the adequate feedback and predictability of the delays resulted in the low stress reported by the participants. The participants knew that the SRT would eventually end and they had constant feedback available during the SRT. The results confirm and strengthen the existing research on SRT and stress (Barron et al., 2004; Boucsein, Baltissen, & Euler, 1984; Kuhman et al. 1990).

Bar Behavior. The comparison of the seven different progress bar behaviors showed that differences exist between the types of behaviors (decelerating, linear, and accelerating), but little difference exist within the three behaviors. Specific progress bar behaviors were optimal for different dependent variables. Each behavior was best on some measure suggesting that a progress bar behavior could be selected depending on the desired effect of the progress bar. When analyzed further, the differences caused by behavior of the progress bar showed clear results.

The ANOVAs performed on this collapsed set of conditions show a clear, significant difference in each of the dependent measures caused by behavior of the progress bar. The linear condition allowed the participants to perform better in the metrics of length of time before switching to task two, number of checks, and length of checks. However, the decelerating progress bar allowed for better participant performance in accuracy of returning to task one for the 10 second duration as well as proportion of late returns. The probable reason for the decelerating progress bar allowing for greater accuracy of return is the increase in resolution of the progress bar as it approaches the end. For example, a linear progress bar will display five seconds of time in the second 50 percent of the progress bar's length, whereas a decelerating progress bar would display as much as eight seconds of time in the second 50 percent of the progress bar's length. Therefore one-fifth of the linear progress bar's remaining 50 percent will equal one second, but one-fifth of the decelerating progress bar's remaining 50 percent would equal as much as 1.6 seconds. By increasing the amount of time the progress bar requires to fill the second half of the bar, the decelerating behavior allows the user to assess the remaining time more accurately.

The second hypothesis that accelerating progress bar should perform worst was supported in only one instance and that the linear progress bar should perform best was supported in three instances: length of time before switching to task one, number of checks on task one, and the length of the checks. The linear progress bar had average performance for accuracy of return to task one and performance on task two. The accelerating progress bar performed worst on accurately returning to task one and it performed best for number of checks and performance on task two. Because the linear progress bar performed best at three of the five metrics and was never the worst, it is concluded that the linear progress bar performed best on average.

The previous research demonstrating that users should favor an accelerating progress bar was not supported (Harrison, et al., 2007). The current results indicate that users preferred a linear progress bar, which is also what they performed best with, which is in agreement with the previous research that performance and preference are linked (Kuhman et al., 1990). In addition, this result conflicts with Conn et al. and their results that show that users prefer accelerating progress bars. Their task may have biased users to prefer the acceleration because all they had to do was watch the progress bar and respond if they preferred it. A second task was not present for them to do, so their entire basis of liking one progress bar over the other is the perceived SRT. The accelerating progress bar is deceptive in making a user believe that time progress is going faster and SRT has been shortened.

The results indicating that checking often leads to a more accurate return to task one is not surprising, but nonetheless interesting. Only for the linear and accelerating behaviors did a strong relationship with number of checks and accuracy of return exist. For these two behaviors, more checks on the progress bar related to a more accurate return to task one. It is surprising though that the decelerating progress bars did not exhibit this relationship as well. It is probable that because the decelerating progress bars had the most accurate returns to begin with, that number of checks did not influence accuracy of return as strongly as it did for the other two behaviors because there was little room for improvement. In general, it appears that participants overestimated the SRT, resulting in late returns to the primary task. The decelerating bar compensated for this bias, resulting in better accuracy.

Duration of SRT. It is common sense that duration of SRT will directly influence the number of completions of the second task that the participant could possibly do. Similarly, a longer SRT will result in more checks on the progress bar because the amount of possible time to check is greater. However, SRT duration did not affect how long the participant took to switch to task two, or how long the checks took. SRT duration affected what it would logically influence, but did not affect anything else.

Interactions. SRT duration by itself did not have an effect on the length of checking back on the progress bar, as did the type of the progress bar, however these two variables together had an effect on the length of checking back on the progress bar. For the shorter duration SRT, the continuous and segmented (M_C = 715 ms, M_S = 733 ms) progress bar conditions were equivalent, but for the longer duration SRT, a significant difference existed between the conditions (M_C = 879 ms, M_S = 960 ms). Although not a large difference between the two means, the implications are important; the longer the duration of the SRT the more difficult it is to get a sense of the time remaining with the segmented progress bar. The interaction between SRT duration and progress bar type for length of checks is surprising, but easily explained. This outcome can be attributed to the design of the segmented bar. For the 10 second SRT, one segment is equal to 1 second, whereas for the 20 second SRT, one segment is equal to 2 seconds. Therefore, to see an equivalent amount of progress—one segment—for the two SRT durations, the user must wait longer for the longer SRT duration. This effect could be prevented by scaling the number of segments to the length of the SRT duration. If 1 segment was always equal to 1 second, doubling the number of segments for the 20 second SRT, the length of check for the two durations would be made equal.

For the accuracy of returning to the primary task at the end of the SRT, both progress bar behavior and duration were significant, as well as the interaction between them. For the 10 second SRT duration, the differences between the three behaviors are rather linear. The decelerating behavior had the lowest error for returning to primary task close to the end of the SRT, the linear behavior was in the middle, and the accelerating behavior was the highest. However, for the 20 second SRT duration, the linear behavior was equal to the decelerating behavior, and the accelerating was the worst. The interaction between progress bar behavior and SRT duration is most prominent for the linear behavior, with a mean error of return for the 10 second SRT duration of 4827 ms and 3080 ms for the 20 second SRT duration, an improvement of nearly 2 seconds.

When considering the past research on motion estimation (Cao, Gu, & Wang, 2004; Caston & Bricout-Berthout, 1985) and time estimation (Wearden and Jones, 2007), the interaction between SRT duration and progress bar behavior has a greater application. SRT duration and progress bar behavior exist at the intersection of these two research topics. The results for the 10 second SRT would suggest that the motion estimation research provides a better theoretical framework for explaining the participants' behavior and the results for the 20 second SRT would suggest that the time estimation research provides a better framework. I theorize that neither one of these research topics alone can solve this problem, and that both are needed to explain the interaction effect.

As a progress bar is a graphic representation of the passage of time, it stands to reason that both time estimation and motion estimation are required to estimate the duration of the SRT that the progress bar represents. The time estimation is dependent on the estimation of the motion of the progress bar, so the motion estimation must occur first and that information be passed to the participant's internal clock. For the short SRT duration, the better performance resulting from the decelerating progress bar behavior is due to a dependence on the estimation of the motion. On average, participants watched 1.5 seconds of the progress bar before switching to task two. For the decelerating progress bar behavior, that would mean that 30% of the progress bar was filled, and for the accelerating behavior, 3% of the progress bar was filled. The motion of the decelerating progress bar would indicate that the end is rapidly approaching causing the participants to return to the primary task sooner than they would for the other progress bar behaviors, which also explains the higher ratio of early to late returns.

For the long SRT duration, the participants again relied on their estimation of time based off of their estimation of the motion of the progress bar. On average, they watched 1.75 seconds of the progress bar before switching to task two. For the decelerating progress bar behavior, that would mean that 18% of the progress bar was filled, and for the accelerating behavior, 6% of the progress bar was filled. If the SRT duration is extrapolated to 30 seconds and the initial viewing period before switching to task two remains constant, 10% of the progress bar will be filled for the decelerating behavior, 5% will be filled for the linear behavior, and 2% for the accelerating behavior. Therefore, as the duration of the SRT increases, the amount of the progress bar that is filled during the initial viewing period will begin to resemble a linear progress bar. The longer SRT duration enables the participant's internal clock to more accurately estimate the duration. The participant's internal clock is being misled by the information it is receiving from the motion estimation mechanism.

Recommendations

The results can be interpreted to indicate that different progress bar speeds should be used depending on the goal of the progress bar. If a user is working on two tasks concurrently and returning on time to the primary task that involves the SRT is more important, the best progress bar is continuous and decelerating. Users are early about half the time with a decelerating progress bar, as opposed to being early only about fifteen percent of the time with an accelerating progress bar. Additionally, when users are late, the amount of time they are late is far less for decelerating bars than accelerating. For the extreme deceleration, participants were late by .35 seconds, whereas for the extreme acceleration participants were late by 5.72 seconds. The cause of this difference arises from the increase in granularity of the progress bar as it approaches the end. By lengthening the amount of time the progress bar requires to fill the second half, the decelerating condition can allow the user to assess the remaining time more accurately.

If vigilance on the first task is most important and the user should spend the most amount of time looking at the primary task, then a decelerating segmented progress bar is best. This combination resulted in the longest amount of time before switching to task two, the most checks on the primary task, and the longest checks, leading the user to spend the most amount of time paying attention to task one. The main cause of the increase of time spent on task one is due to the nature of the segmented bar, as discussed earlier. For the user to gain any information from the segmented bar, the user must witness one segment appearing. To gain knowledge on the behavior of the progress bar, the user must observe at least two segments being completed. From the two segments it can be determined if the progress bar is accelerating or decelerating.

If performance on the second task is more important, then a continuous and accelerating progress bar is best. This combination resulted in the most completions of the second task, on average three more completions than the decelerating and continuous progress bar. This is most likely due to the user's initial perception of the progress bar. The slowly filling progress bar could indicate to the user that the SRT is going to be quite lengthy, so a lot of time can be spent on the second task. The belief that the SRT will be long also leads the user to be late a majority of the time (Fig 1). The opposite of this may also explain why the decelerating progress bars results in fewer late returns, the progress bar begins filling rapidly leading the user to believe that the SRT is very short, which in turn results in being early more often.

By comparing the effects on performance due to the continuous and segmented progress bars, an additional item can be added to the list of what is required for a good time affordance developed by Conn (1995). This additional item would be the resolution of the progress bar, or what minimum amount of time is needed to perceive a change in the progress bar. The continuous progress bar would have the greatest amount of resolution as it is constantly in motion, which would be in effect a segmented progress bar with an infinite amount of segments. Depending on the application of the progress bar, different levels of resolution may have advantages; for instance, if vigilance on the primary task is most important, as recommended earlier. It is the author's recommendation that resolution should be added to the list of criteria for a good time affordance.

Conclusion

In conclusion, participants performed best with the continuous progress bar as opposed to the segmented progress bar and preferred the continuous progress bar. Second, participants performed best—on average—with the linear progress bar, and also favored this behavior. And third, performance was affected by the duration of the SRT, with longer SRT allowing for the user to work longer on the secondary task and checking back on the progress bar more. This study looked at progress bars that behaved in a constant, predictable motion, which is seldom how actual progress bars behave in a system that is influenced by external factors. When external factors are not at play and the SRT is predictable and reliable, the recommendations of this study can be applied.

In actual computer use, progress bars seldom follow a single behavior. Often, they speed up and slow down – seemingly under their own will – at different times throughout the duration of the SRT. Previous research has shown that users do not favor these kinds of progress bars, but they are common. Therefore, an additional topic that that warrants further exploration is mixed speed progress bars. Research should explore how these affect performance on primary and secondary tasks. A progress bar that behaves linearly for the first half and then decelerates for the second half may allow the user to comprehend how long the SRT will be from the linear portion and estimate the end accurately from the decelerating portion. A second influence on how late or early a participant is to return to the primary task is the number of checks on the SRT the participant makes. A user who checks more often is more likely to return to the primary task accurately.

Future research should investigate the thought process of a user checking on the primary task. An interesting experiment would be to look at the difference between checking on the primary task and returning to the second task or checking on the primary task and then waiting for it to end. This question would answer at what point is “too soon” to return and wait, or when the user decides that they can get more work done on another task.

From the results of the current study, the accelerating progress bar allows for success at the second task and the linear progress bar allows for success at the primary task. Perhaps a progress bar that follows both of these behaviors would allow for success at both tasks. A bar that accelerates up to the mid point and then behaves linearly to the end might facilitate performance on multiple tasks.

References

Apple Inc., (2006). *Apple human interface guidelines* [pp. 308-313]. (Adobe PDF), Retrieved from http://developer.apple.com/mac/library/documentation/UserExperience /Conceptual/AppleHIGuidelines/XHIGIntro/XHIGIntro.html

Barron, K., Simpson, T. W., Rothrock, L., Frecker, M., Barton, R. R., & Ligetti, C. (2004). *Graphical User Interfaces for Engineering Design: Impact of Response Delay and Training on User Performance.* Paper presented at the Proceedings of DETC'04 ASME 2004 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Salt Lake City, Utah, USA.

Boucsein, W., Baltissen, R., & Euler, R. (1984). Dependence of skin-conductance reactions and skin resistance reactions upon previous level. *Psychophysiology, 21*(2), 212-218.

Brouwer, A., Brenner, E., & Smeets, J. (2002). Perception of acceleration with short presentation times: can acceleration be used in inception? *Perception & Psychophysics*, 64, 1160–1168.

Cao, P., Gu, Y., & Wang, S.R., (2004). Visual neurons in the pigeon brain encode the acceleration of stimulus motion. *The Journal of Neuroscience*, *24*(35), 7690-7698

Caston, J., & Bricout-Berthout, A. (1958). Influence of stimulation of the visual system on the activity of vestibular nuclear neurons in the frog. *Brain, Behavior, and Evolution, 26*, 49-57.

Carbonell J.R., Elkind, J.I., & Nickerson, R.S. (1968). On the psychological importance of time in a time sharing system. *Human Factors, 10*(2), 135-142.

Conn, A. P. (1995). *Time Affordances: The Time Factor in Diagnostic Usability Heuristics.* Paper presented at the Conference on Human Factors in Computing Systems, Denver, CO, USA.

Coren S., Girgus J.S. (1978) *Seeing is Deceiving: The Psychology of Visual Illusions*. Lawrence Erlbaum Associates, New Jersey.

Craig, J. C. (1973). A constant error in the perception of brief temporal intervals. *Perception & Psychophysics, 13*, 99-104.

Dannenbring, G. L. (1983). The effect of computer-response time on user performance and satisfaction - a preliminary investigation. *Behavior Research Methods & Instrumentation, 15*(2), 213-216.

DeLucia, P. R., & Liddell, G. W. (1998). Cognitive motion extrapolation and cognitive clocking in prediction motion tasks. *Journal of Experimental Psychology: Human Perception and Performance, 24*(3), 901–914.

Eggemeier, F., Wilson, G., Kramer, A., & Damos, D. (1991). Workload assessment in multi-task environments. In D. Damos (Ed) *Multiple Task Performance*. 207-216, Washington, DC: Taylor & Francis Ltd.

Harrison, C., Amento, B., Kuznetsov, S., & Bell, R. (2007). *Rethinking the Progress Bar.* Paper presented at the Proceedings of the 20th Annual ACM Symposium on User interface Software and Technology, Newport, RI, USA.

Hart, S.G., & Staveland, L. (1988). Development of the NASA task load index (TLX): results of empirical and theoretical research. In P.A. Hancock and N. Meshkati (Eds.) *Human Mental Workload*, 139-183. Amsterdam: North- Holland.

Hendy, K. C., Hamilton, K. M. Landry, L. N. (1993). Measuring subjective workload: When is one scale better than many. *Human Factors, 35*(4), 579-601.

Hitt, J. M., Kring, J. P.,Daskarolis, E., Morris, C., Mouloua, M. (1999). *Assessing mental workload with subjective measures: An analytical review of the NASA-TLX index since its inception.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting. Santa Monica, CA.

Kuhmann, W., Boucsein, W., Schaefer, F., & Alexander, J. (1987). Experimental investigation of psychophysiological stress-reactions induced by different system response-times in human-computer interaction. *Ergonomics, 30*(6), 933-943.

Kuhmann, W., Schaefer, F., & Boucsein, W. (1990). Effects of waiting time in simple task performance-analogy to computer-response times in man-machine interaction. *Zeitschrift Fur Experimentelle Und Angewandte Psychologie, 37*(2), 242-265.

Lightner, N. J., Bose, I., & Salvendy, G. (1996). What is wrong with the world-wide web. *Ergonomics, 39*(8), 996-1004.

Lyon, D. R., & Waag, W. L. (1995). Time-course of visual extrapolation accuracy. *Acta Psychologica, 89*(3), 239–260.

Mace, M. (2007). Growth of web applications in the US: rapid adoption, but only when there's a real benefit. Retrieved January 10, 2008, from http://www.rubiconconsulting .com /thinking/whitepaper/ 2007/09/growth_of_web_applications_in.html.

Martin, G., & Corl, K. (1986). System response time effects on user productivity. *Behavior & Information Technology, 5*, 3-13.

Meyer, J., Bitan, Y., & Shinar, D. (1995). Displaying a boundary in graphic and symbolic "Wait" displays: duration estimates and users' preferences. *International Journal of Human-Computer Interaction, 7*(3), 273-290.

Meyer, J., Shinar, D., Bitan, Y., & Leiser, D. (1996). Duration estimates and users' preferences in human-computer interaction. *Ergonomics, 39*(1), 46-60.

Miller, R.B. (1968) Response Time in man-computer conversational transactions. In *Proc Fall Joint Computer Conference. AFIPS*. San Francisco. CA. 267-277.

Myer, H. A., & Hildebrandt, M. (2002). *Towards time design: pacing of hypertext navigation by system response times*. Paper presented at the CHI '02 extended abstracts on Human factors in computing systems.

Myers, B. A. (1985). The importance of percent-done progress indicators for computer-human interfaces. *Proc. ACM CHI'85 Conf. (San Francisco, CA, 14-18 April)*, 11-17.

Morfield. M.A., Wiesen. R.A., Grossberg, M., and Yntema, D.B. (1969) Initial experiments on the effects of system delay on on-line problem solving. *Lincoln Laboratory Tech. Report.*

Nielson, J. (1994). *Usability Engineering*. San Francisco, CA: Morgan Kaufmann.

Ornstein, R. E. (1969). *On the Experience of Time*. Baltimore, MD: Penguin Books.

Port, N. L., Lee, D., Dassonville, P., & Georgopoulos, A. P. (1997). Manual interception of moving targets: I. Performance and movement initiation. *Experimental Brain Research*, 116, 406–420.

Poynter, W. D. (1989). Judging the duration of time intervals: a process of remembering segments of experience. In I. Levin, and Zakay, D. (Ed.), *Time and Human Cognition*. Amsterdam: Elsevier and North-Holland.

Rosenbaum, D.A. (1975). Perception and extrapolation of velocity and acceleration. *Journal of Experimental Psychology: Human Perception and Performance, 1*(4), 395-403.

Rushinek, A., & Rushinek, S. F. (1986). What makes users happy - the results of a user questionnaire are used to determine the effects of 17 different independent variables on user satisfaction. *Communications of the Acm, 29*(7), 594-598.

Schaefer, F. (1990). The effect of system response times on temporal predictability of work flow in human-computer interaction. *Human Performance, 3*(3), 173-186.

Schleifer, L. M., & Amick, B. C. (1989). System response time and method of pay: stress effects in computer-based tasks. *International Journal of Human-Computer Interaction, 1*(1), 23-39.

Selvidge, P. R., Chaparro, B. S., & Bender, G. T. (2001). The world-wide wait: effects of delays on user performance. *International Journal of Industrial Ergonomics, 29*, 15-20.

Senot, P., Prévost, P., McIntyre, J. (2003). Estimating time to contact and impact velocity when catching an accelerating object with the hand. *Journal of Experimental Psychology: Human Perception and Performance, 29*(1), 219-237.

Shneiderman, B. (1987). *Designing the User Interface*. New York, NY: Addison Wesley.

Steiner, S. (1968). Apparent duration of auditory stimuli. *Journal of Auditor Research, 8*, 195-205.

Thum, M., Boucsein, W., Kuhman, W., & Ray, W. J. (1995). Standardized task strain and system response times in human computer interaction. *Ergonomics, 38*(7), 1342-1351.

Tresilian, J. R. (1995). Perceptual and cognitive-processes in time-to-contact estimation: Analysis of prediction-motion and relative judgment tasks. *Perception & Psychophysics, 57*(2), 231–245.

Wearden, J. H., Denovan, L., Fakhri, M., & Haworth, R. (1997). Scalar timing in temporal generalization in humans with longer stimulus durations. Journal of Experimental *Psychology: Animal Behavior Processes, 23*, 502–511.

Processes, 23, 502–511.Wearden, J. H., & Jones, L. A. (2007) Is the growth of subjective time in humans a linear or nonlinear function of real time? *The Quarterly Journal of Experimental Psychology, 60*(9), 1289–1302.

Wearden, J. H., & McShane, B. (1988). Interval production as an analogue of the peak procedure: Evidence for similarity of human and animal timing processes. *Quarterly Journal of Experimental Psychology, 40B*, 363–375.

Werkhoven, P., Snippe, H. P., & Toet, A. (1992). Visual processing of optic acceleration. *Vision Research*, 32, 2313–2329.

Appendix A

Post-Section Questionnaire

1. How mentally demanding was the task?

Very Low	1	2	3	4	5	6	7	Very High

2. How hurried or rushed was the pace of the task?

Very Low	1	2	3	4	5	6	7	Very High

3. How successful were you in accomplishing what you were asked to do?

Perfect	1	2	3	4	5	6	7	Failure

4. How hard did you have to work to accomplish your level of performance?

Very Low	1	2	3	4	5	6	7	Very High

5. How insecure, discouraged, irritated, stressed, and annoyed were you?

Very Low	1	2	3	4	5	6	7	Very High

6. I thought the progress indicator did a satisfactory job of representing the time delay

Agree	1	2	3	4	5	6	7	Disagree

7. How long did you feel the searches took to complete?

_________ Seconds

Post-Test Questionnaire

1. Which progress indicator did you prefer?
   - ☐ Solid bar
   - ☐ Blocked bar

2. Which progress indicator behavior did you prefer?
   - ☐ Accelerating
   - ☐ Constant
   - ☐ Decelerating

3. How often do you use desktop or laptop computers?
   - ☐ Once a week
   - ☐ A day or two a week
   - ☐ Every day
   - ☐ Several times a day

4. How much experience do you have with installing programs on a computer?
   - ☐ No experience
   - ☐ A little experience
   - ☐ Some experience
   - ☐ A lot of experience

5. How often do you download files from the internet, or transfer files between computers?
   - ☐ Once a week
   - ☐ A day or two a week
   - ☐ Every day
   - ☐ Several times a day

Sex: M F

Age: ________

Appendix B

List of search terms and system responses

	search term	response		search term	response
1	dishonesty	cheating	**43**	blowgun	weapon
2	slightly	somewhat	**44**	sportsman	athlete
3	stupefy	confuse	**45**	precarious	unstable
4	superhuman	heroic	**46**	variable	uncertain
5	freely	willfully	**47**	rain	precipitation
6	impudence	boldness	**48**	athletic	fit
7	acceptable	passable	**49**	ruinous	misfortunate
8	authorized	sanctioned	**50**	mallet	hammer
9	dismantle	level	**51**	turf	sod
10	engraving	print	**52**	cavalcade	fleet
11	vegetate	idle	**53**	aloud	audibly
12	weakly	delicately	**54**	atheistic	nonbelief
13	affliction	malady	**55**	emancipate	free
14	behavior	conduct	**56**	anecdote	quote
15	indefinite	vague	**57**	lacy	frilly
16	vigilance	watchfulness	**58**	disposition	mood
17	violation	crime	**59**	frighten	alarm
18	frustrate	torment	**60**	artificiality	falsehood
19	acknowledgement	admission	**61**	pacific	oceanic
20	unschooled	untaught	**62**	mart	store
21	darken	dull	**63**	funeral	burial
22	inherited	inborn	**64**	undulation	wave
23	nurture	foster	**65**	stumpy	short
24	correlation	relationship	**66**	repercussion	consequence
25	crooked	misleading	**67**	subcommittee	group
26	cleanse	wash	**68**	hygiene	health
27	creative	original	**69**	interweave	entwine
28	evacuate	flee	**70**	oxen	cattle
29	ineffable	indefinable	**71**	allotment	amount
30	archives	documents	**72**	redden	blush
31	shipyard	port	**73**	conserve	save
32	rerun	replay	**74**	kneel	submit
33	crook	criminal	**75**	public	free
34	demagnetize	allure	**76**	competent	intelligent
35	insurgence	rebellion	**77**	aseptic	infected
36	recondition	renovate	**78**	stationery	paper
37	torso	chest	**79**	information	news
38	abduct	capture	**80**	bookish	studious
39	feral	domestic	**81**	water	aqua
40	nature	character	**82**	platform	stand
41	nation	country	**83**	propeller	prop
42	ice	snow	**84**	sphere	globe

Appendix C

Key for determining condition shown in figures:

CON = Continuous progress bar

SEG = Segmented progress bar

A3 = Accelerating rate of change 3

A2 = Accelerating rate of change 2

A1 = Accelerating rate of change 1

L = Linear rate of change

D1 = Decelerating rate of change 1

D2 = Decelerating rate of change 2

D3 = Decelerating rate of change 3

10 = 10 SRT duration

20 = 20 SRT duration

Figure C1 Histogram of length of time before first switch to task two in ms per condition

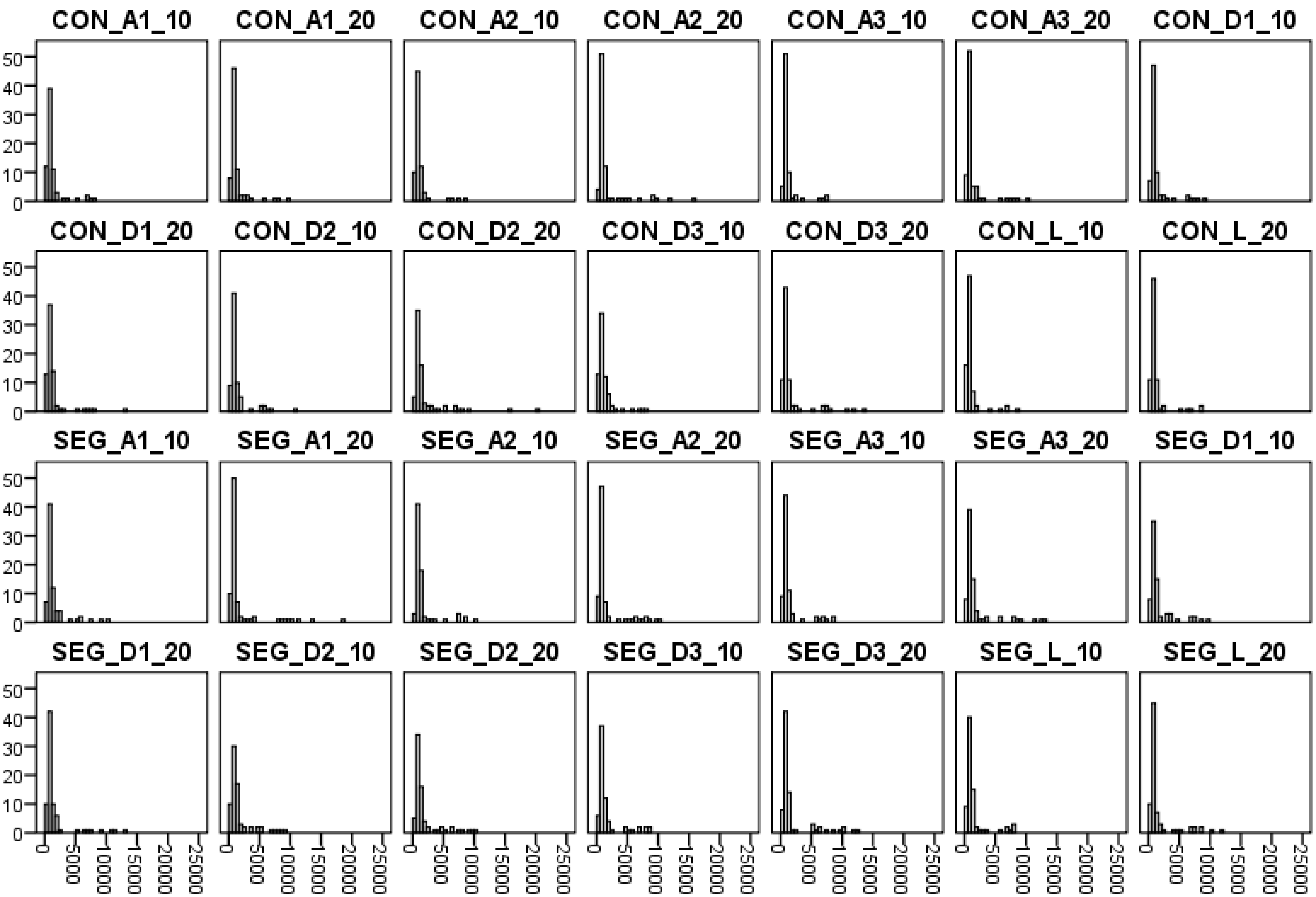


Figure C2 Histogram of number of checks on task one per condition

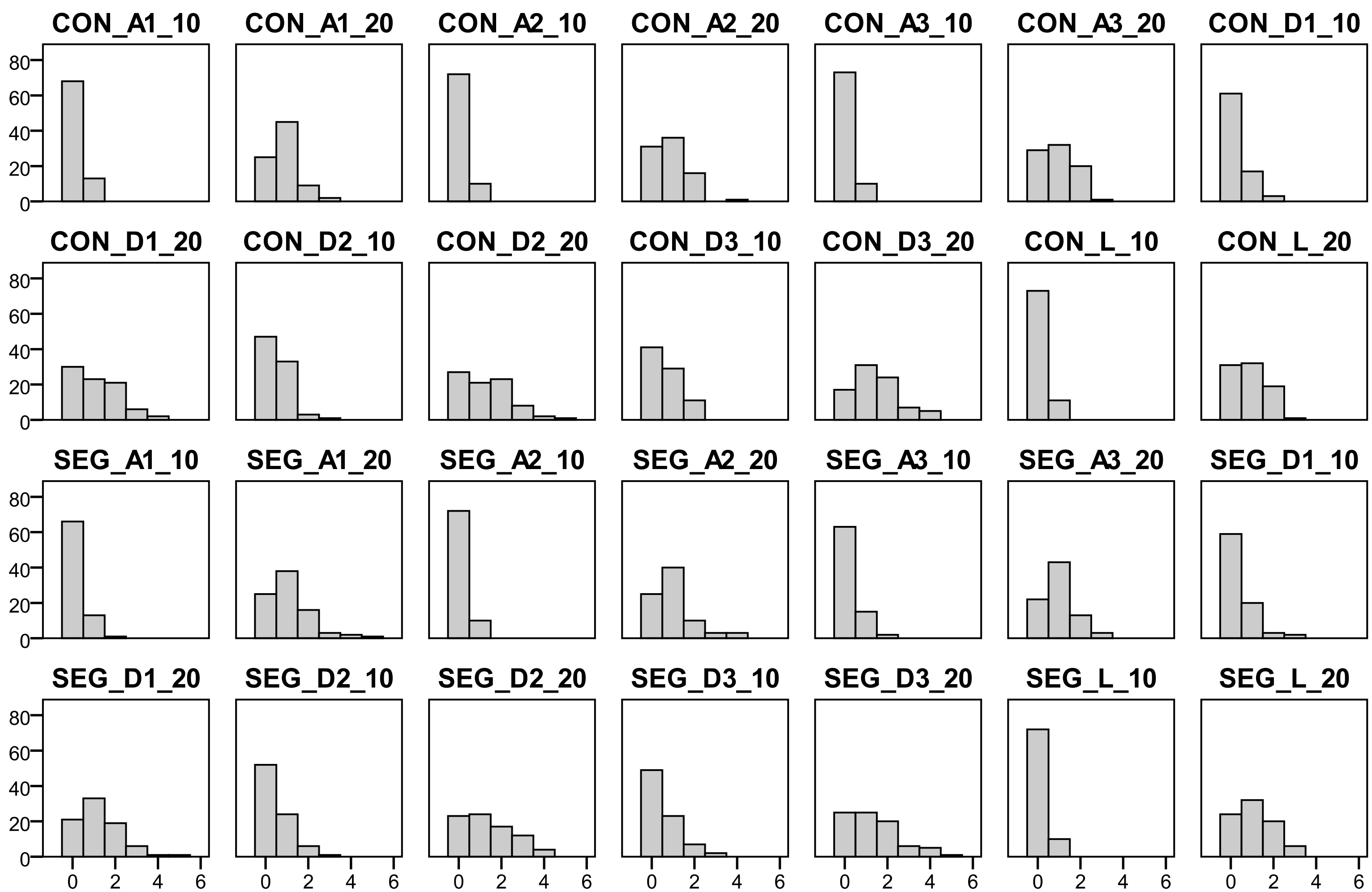


Figure C3 Histogram of length of checks on progress bar in ms per condition

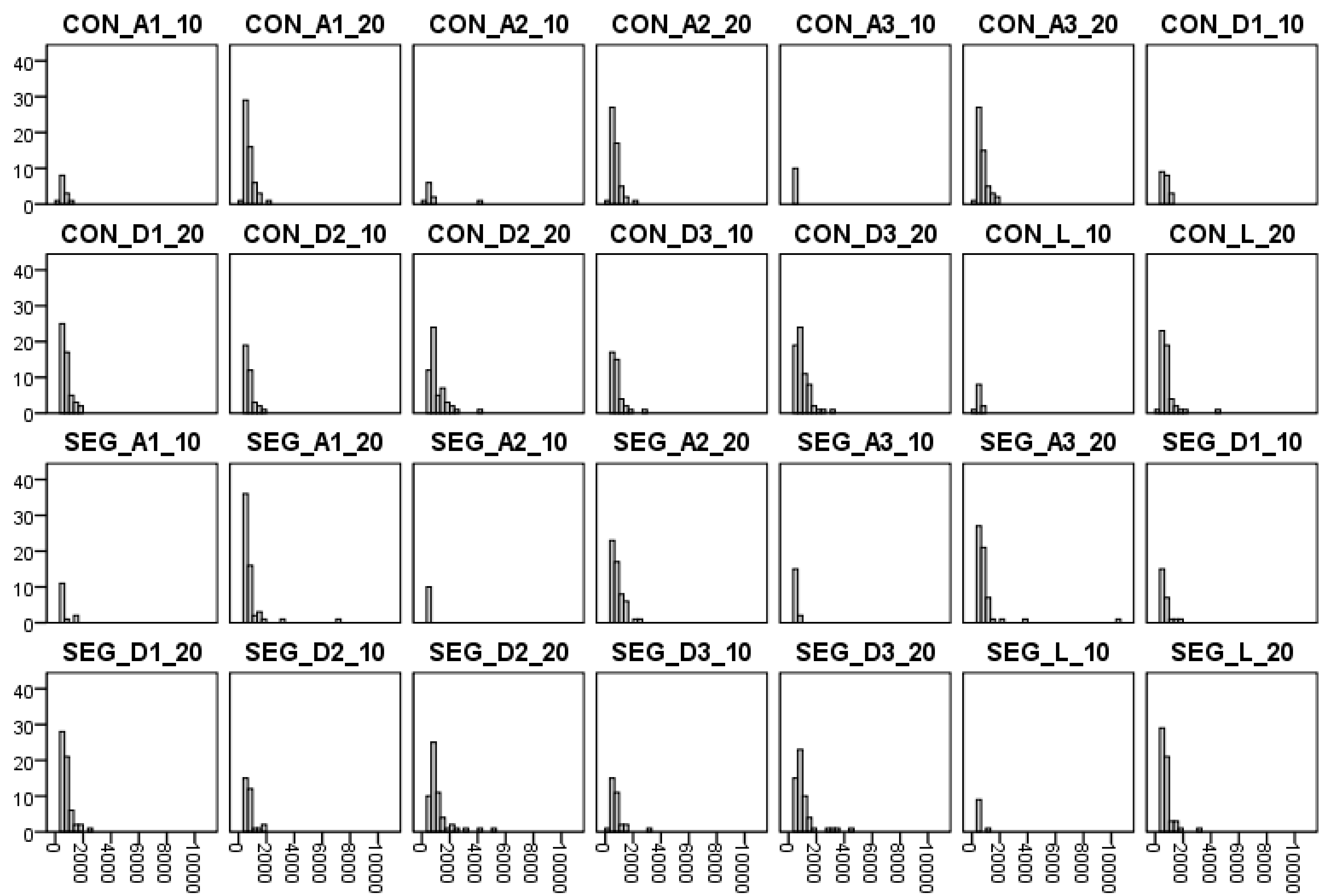


Figure C4 Histogram of accuracy of return to task one in ms per condition

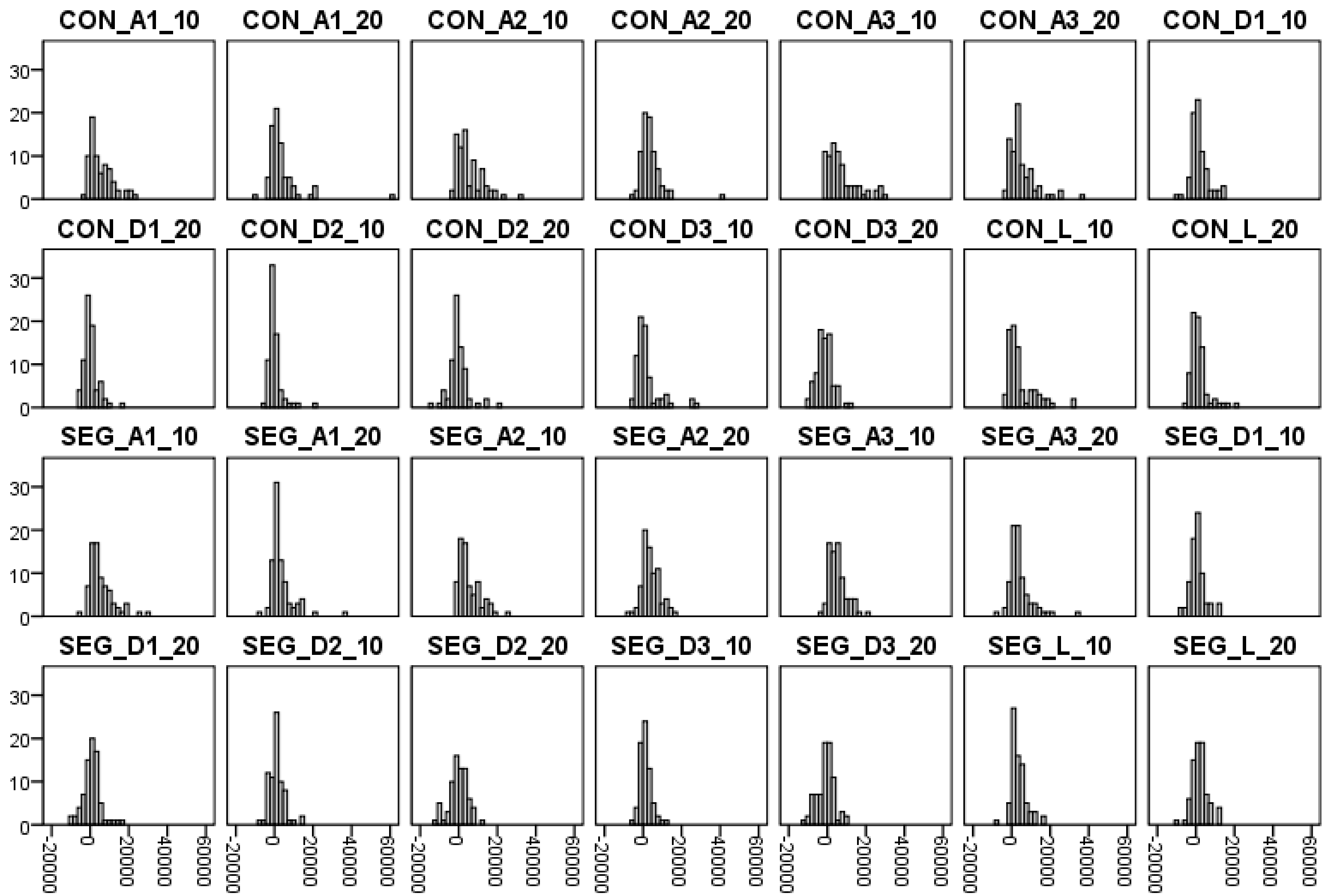
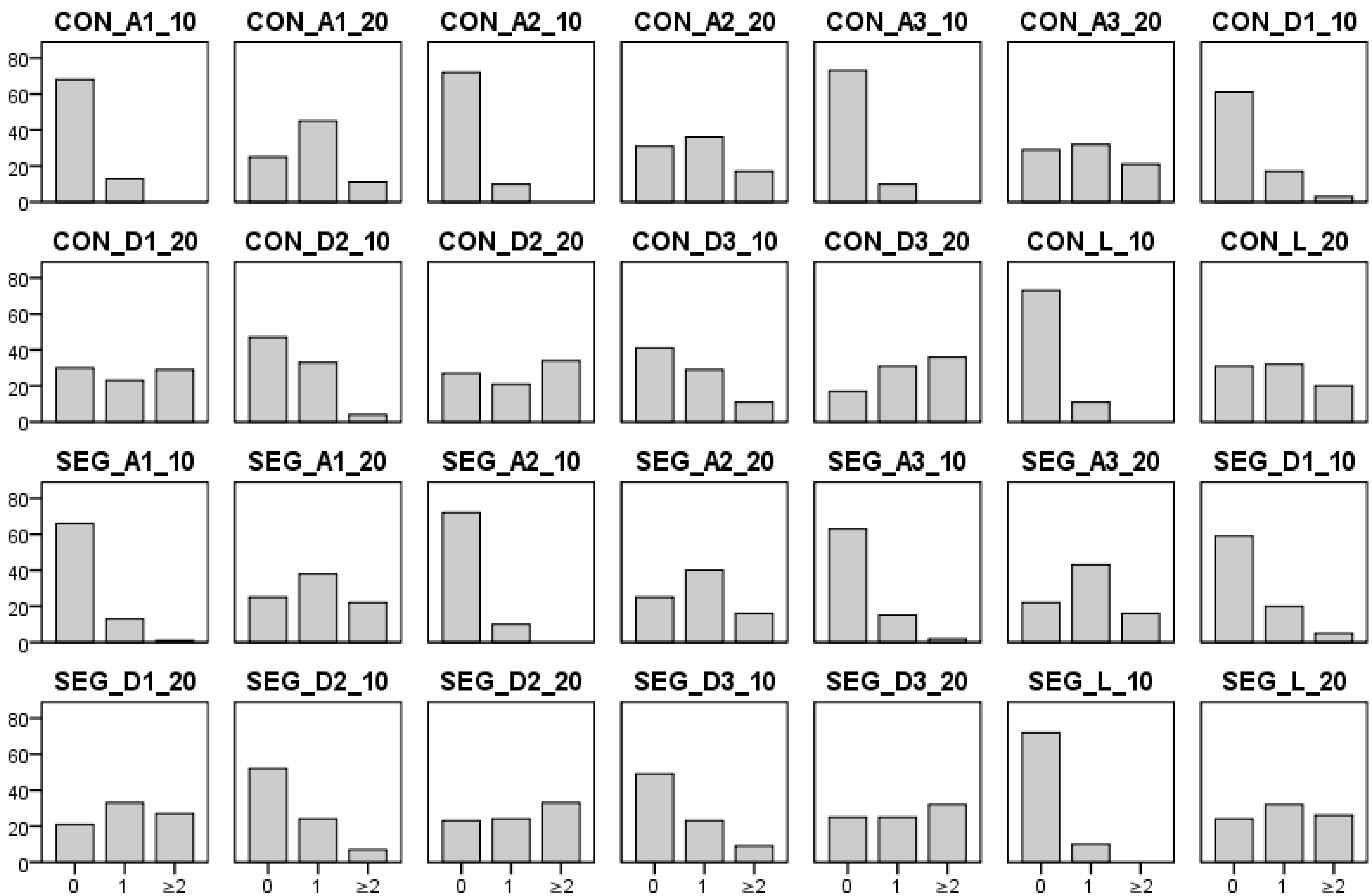


Figure C5 Histogram of log transform of length of time before switching to task two per condition



Appendix D

Table D1 Repeated measures ANOVA of length of time before switching to task two

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	5008.545	1	5008.545	271852.217	0.000
Subject	28.860	26	1.110	60.249	0.000
Type	0.241	1	0.241	13.074	0.000
Behavior	0.318	6	0.053	2.874	0.009
Duration	0.020	1	0.020	1.083	0.298
Type * Behavior	0.051	6	0.008	0.461	0.838
Behavior * Duration	0.058	6	0.010	0.526	0.789
Type * Duration	0.008	1	0.008	0.407	0.524
Type * Behavior * Duration	0.192	6	0.032	1.733	0.111
Error	10.115	549	0.018		

Table D2 Repeated measures ANOVA of length of checks on task one

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	3.181E+07	1	3.181E+07	790.982	0.000
Subject	8.647E+06	21	4.118E+05	10.237	0.000
Type	2.513E+04	1	2.513E+04	0.625	0.431
Behavior	9.822E+05	6	1.637E+05	4.070	0.001
Duration	9.025E+04	1	9.025E+04	2.244	0.137
Type * Behavior	1.967E+05	6	3.278E+04	0.815	0.560
Behavior * Duration	3.673E+05	6	6.122E+04	1.522	0.178
Type * Duration	672.539	1	672.539	0.017	0.897
Type * Behavior * Duration	5.368E+05	4	1.342E+05	3.337	0.013
Error	4.344E+06	108	4.022E+04		

Table D3 Repeated measures ANOVA of accuracy of return to task one

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	5.424E+09	1	5.424E+09	708.130	0.000
Subject	5.017E+09	26	1.930E+08	25.191	0.000
Type	947.660	1	947.660	0.000	0.991
Behavior	1.973E+09	6	3.288E+08	42.926	0.000
Duration	2.619E+08	1	2.619E+08	34.186	0.000
Type * Behavior	5.413E+07	6	9.022E+06	1.178	0.316
Behavior * Duration	3.066E+07	6	5.111E+06	0.667	0.676
Type * Duration	1.624E+07	1	1.624E+07	2.120	0.146
Type * Behavior * Duration	2.082E+07	6	3.470E+06	0.453	0.843
Error	4.220E+09	551	7.660E+06		

Table D4 Repeated measures ANOVA of number of completions of task two

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	127636.572	1	127636.572	6720.610	0.000
Subject	10164.813	26	390.954	20.585	0.000
Type	13.600	1	13.600	0.716	0.398
Behavior	3798.641	6	633.107	33.336	0.000
Duration	3078.594	1	3078.594	162.101	0.000
Type * Behavior	29.773	1	29.773	1.568	0.211
Behavior * Duration	47.934	6	7.989	0.421	0.865
Type * Duration	94.934	6	15.764	0.830	0.547
Type * Behavior * Duration	84.762	6	14.127	0.744	0.614
Error	12800.482	674	18.992		

Table D5 Repeated measures ANOVA of number of checks on task one

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	342.897	1	342.897	1865.037	0.000
Subject	139.223	26	5.355	29.125	0.000
Type	1.817	1	1.187	9.884	0.002
Behavior	18.694	6	3.116	16.947	0.000
Duration	44.670	1	44.670	242.966	0.000
Type * Behavior	0.997	6	0.166	0.904	0.491
Behavior * Duration	1.275	6	0.213	1.156	0.328
Type * Duration	0.067	1	0.067	0.366	0.545
Type * Behavior * Duration	0.921	6	0.153	0.835	0.543
Error	123.918	674	0.184		

Appendix E

Table E1 Repeated measures ANOVA of length of time before switching to task two with progress bar behavior averaged

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	2426.438	1	2426.438	212106.922	.000
Subject	15.275	26	.588	51.357	.000
Type	.155	1	.155	13.536	.000
Behavior	.188	2	.094	8.214	.000
Duration	.021	1	.021	1.869	.173
Type * Behavior	.009	2	.004	.390	.678
Duration * Behavior	.002	2	.001	.092	.912
Duration * Type	.015	1	.015	1.299	.256
Duration * Type * Behavior	.004	2	.002	.191	.826
Error	2.791	244	.011		

Table E2 Repeated measures ANOVA of number of checks on task one with behavior collapsed

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	134.898	1	134.898	1033.822	.000
Subject	53.674	26	2.064	15.821	.000
Type	1.053	1	1.053	8.074	.005
Behavior	7.218	2	3.609	27.658	.000
Duration	51.786	1	51.786	396.875	.000
Type * Behavior	.103	2	.051	.393	.675
Duration * Behavior	.088	2	.044	.336	.715
Duration * Type	.412	1	.412	3.155	.077
Duration * Type * Behavior	.191	2	.095	.731	.482
Error	35.753	274	.130		

Table E3 Repeated measures ANOVA of length of checks on task one with progress bar behavior averaged

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	21885867.515	1	21885867.515	206.971	.000
Subject	11318382.835	25	452735.313	4.281	.000
Type	7897.408	1	7897.408	.075	.785
Behavior	240973.793	2	120486.896	1.139	.324
Duration	39008.969	1	39008.969	.369	.545
Type * Behavior	254606.456	2	127303.228	1.204	.305
Duration * Behavior	134816.928	2	67408.464	.637	.531
Duration * Type	623721.820	1	623721.820	5.898	.017
Duration * Type * Behavior	257062.522	1	257062.522	2.431	.122
Error	9939883.441	94	105743.441		

Table E4 Repeated measures ANOVA of accuracy of return to task one with progress bar behavior averaged

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	5.217E9	1	5.217E9	928.458	.000
Subject	1.731E9	26	66580759.120	11.850	.000
Type	11835746.767	1	11835746.767	2.106	.148
Behavior	2.682E8	2	1.341E8	23.870	.000
Duration	64834560.104	1	64834560.104	11.539	.001
Type * Behavior	10772481.341	2	5386240.670	.959	.385
Duration * Behavior	70255536.918	2	35127768.459	6.252	.002
Duration * Type	15600926.078	1	15600926.078	2.777	.097
Duration * Type * Behavior	5089547.384	2	2544773.692	.453	.636
Error	1.371E9	244	5618775.457		

Table E5 Repeated measures ANOVA of number of completions of task two with progress bar behavior averaged

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	18583.711	1	18583.711	4079.974	.000
Subject	1545.312	26	59.435	13.049	.000
Type	5.896	1	5.896	1.294	.256
Behavior	448.531	2	224.266	49.237	.000
Duration	953.751	1	953.751	209.392	.000
Type * Behavior	3.193	2	1.597	.351	.705
Duration * Behavior	22.091	2	11.045	2.425	.090
Duration * Type	3.225	1	3.225	.708	.401
Duration * Type * Behavior	.488	2	.244	.054	.948
Error	1248.032	274	4.555		

Appendix F

Table F1 Tukey HSD table of log transform of mean length of time before first switch to task two in ms

Behavior (I)	Behavior (J)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
D3	D2	-0.020	.02130	.962	-.0835	.0426
	D1	-0.001	.02076	1.000	-.0621	.0608
	L	0.049	.02058	.210	-.0120	.1098
	A1	0.028	.02070	.831	-.0334	.0891
	A2	0.022	.02058	.931	-.0385	.0833
	A3	0.008	.02031	1.000	-.0520	.0682
D2	D3	0.020	.02130	.962	-.0426	.0835
	D1	0.020	.02135	.968	-.0434	.0830
	L	0.0694*	.02118	.019	.0067	.1320
	A1	0.048	.02130	.262	-.0148	.1113
	A2	0.043	.02118	.400	-.0198	.1056
	A3	0.029	.02091	.820	-.0333	.0905
D1	D3	0.001	.02076	1.000	-.0608	.0621
	D2	-0.020	.02135	.968	-.0830	.0434
	L	0.050	.02064	.200	-.0115	.1107
	A1	0.028	.02076	.817	-.0330	.0899
	A2	0.023	.02064	.922	-.0380	.0842
	A3	0.009	.02037	1.000	-.0515	.0691
L	D3	-0.049	.02058	.210	-.1098	.0120
	D2	-0.0694*	.02118	.019	-.1320	-.0067
	D1	-0.050	.02064	.200	-.1107	.0115
	A1	-0.021	.02058	.948	-.0820	.0398
	A2	-0.026	.02046	.855	-.0870	.0341
	A3	-0.041	.02019	.403	-.1005	.0189
A1	D3	-0.028	.02070	.831	-.0891	.0334
	D2	-0.048	.02130	.262	-.1113	.0148
	D1	-0.028	.02076	.817	-.0899	.0330
	L	0.021	.02058	.948	-.0398	.0820
	A2	-0.005	.02058	1.000	-.0663	.0555
	A3	-0.020	.02031	.960	-.0798	.0404
A2	D3	-0.022	.02058	.931	-.0833	.0385
	D2	-0.043	.02118	.400	-.1056	.0198
	D1	-0.023	.02064	.922	-.0842	.0380
	L	0.026	.02046	.855	-.0341	.0870
	A1	0.005	.02058	1.000	-.0555	.0663
	A3	-0.014	.02019	.992	-.0740	.0454
A3	D3	-0.008	.02031	1.000	-.0682	.0520
	D2	-0.029	.02091	.820	-.0905	.0333
	D1	-0.009	.02037	1.000	-.0691	.0515
	L	0.041	.02019	.403	-.0189	.1005
	A1	0.020	.02031	.960	-.0404	.0798
	A2	0.014	.02019	.992	-.0454	.0740

Based on observed means. The error term is Mean Square(Error) = .018.

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

Table F2 Tukey HSD table of mean number of checks on the progress bar per trial

Behavior (I)	Behavior (J)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
D3	D2	.09	.059	.771	-.09	.26
	D1	0.21*	.059	.006	.04	.39
	L	0.38*	.059	.000	.20	.55
	A1	0.42*	.059	.000	.24	.60
	A2	0.41*	.059	.000	.24	.59
	A3	0.39*	.059	.000	.22	.57
D2	D3	-.09	.059	.771	-.26	.09
	D1	.13	.059	.321	-.05	.30
	L	0.29*	.059	.000	.12	.47
	A1	0.33*	.059	.000	.16	.51
	A2	0.33*	.059	.000	.15	.50
	A3	0.31*	.059	.000	.13	.48
D1	D3	-0.21*	.059	.006	-.39	-.04
	D2	-.13	.059	.321	-.30	.05
	L	.16	.059	.088	-.01	.34
	A1	0.21*	.059	.011	.03	.38
	A2	0.2*	.059	.015	.02	.37
	A3	0.18*	.059	.042	.00	.36
L	D3	-0.38*	.059	.000	-.55	-.20
	D2	-0.29*	.059	.000	-.47	-.12
	D1	-.16	.059	.088	-.34	.01
	A1	.04	.059	.993	-.13	.22
	A2	.04	.059	.997	-.14	.21
	A3	.02	.059	1.000	-.16	.19
A1	D3	-0.42*	.059	.000	-.60	-.24
	D2	-0.33*	.059	.000	-.51	-.16
	D1	-0.21*	.059	.011	-.38	-.03
	L	-.04	.059	.993	-.22	.13
	A2	-.01	.059	1.000	-.18	.17
	A3	-.03	.059	1.000	-.20	.15
A2	D3	-0.41*	.059	.000	-.59	-.24
	D2	-0.33*	.059	.000	-.50	-.15
	D1	-0.2*	.059	.015	-.37	-.02
	L	-.04	.059	.997	-.21	.14
	A1	.01	.059	1.000	-.17	.18
	A3	-.02	.059	1.000	-.20	.16
A3	D3	-0.39*	.059	.000	-.57	-.22
	D2	-0.31*	.059	.000	-.48	-.13
	D1	-0.18*	.059	.042	-.36	.00
	L	-.02	.059	1.000	-.19	.16
	A1	.03	.059	1.000	-.15	.20
	A2	.02	.059	1.000	-.16	.20

Based on observed means. The error term is Mean Square(Error) = .184.

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

Table F3 Tukey HSD table of mean length of check on progress bar in ms

Behavior (I)	Behavior (J)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
D3	D2	-65.80	51.783	.864	-221.39	89.79
	D1	142.41	56.294	.159	-26.73	311.55
	L	186.54	62.085	.050	-.01	373.08
	A1	159.11	63.421	.167	-31.45	349.66
	A2	172.92	57.895	.052	-1.03	346.87
	A3	220.44*	55.583	.002	53.43	387.45
D2	D3	65.80	51.783	.864	-89.79	221.39
	D1	208.21*	56.294	.006	39.06	377.35
	L	252.34*	62.085	.002	65.79	438.88
	A1	224.91*	63.421	.010	34.35	415.46
	A2	238.72*	57.895	.001	64.76	412.67
	A3	286.24*	55.583	.000	119.23	453.25
D1	D3	-142.41	56.294	.159	-311.55	26.73
	D2	-208.21*	56.294	.006	-377.35	-39.06
	L	44.13	65.895	.994	-153.86	242.12
	A1	16.70	67.155	1.000	-185.08	218.47
	A2	30.51	61.963	.999	-155.67	216.69
	A3	78.03	59.808	.848	-101.67	257.73
L	D3	-186.54	62.085	.050	-373.08	.01
	D2	-252.34*	62.085	.002	-438.88	-65.79
	D1	-44.13	65.895	.994	-242.12	153.86
	A1	-27.43	72.079	1.000	-244.00	189.14
	A2	-13.62	67.268	1.000	-215.73	188.50
	A3	33.90	65.289	.999	-162.27	230.07
A1	D3	-159.11	63.421	.167	-349.66	31.45
	D2	-224.91*	63.421	.010	-415.46	-34.35
	D1	-16.70	67.155	1.000	-218.47	185.08
	L	27.43	72.079	1.000	-189.14	244.00
	A2	13.81	68.502	1.000	-192.01	219.64
	A3	61.33	66.560	.968	-138.66	261.32
A2	D3	-172.92	57.895	.052	-346.87	1.03
	D2	-238.72*	57.895	.001	-412.67	-64.76
	D1	-30.51	61.963	.999	-216.69	155.67
	L	13.62	67.268	1.000	-188.50	215.73
	A1	-13.81	68.502	1.000	-219.64	192.01
	A3	47.52	61.318	.987	-136.72	231.76
A3	D3	-220.44*	55.583	.002	-387.45	-53.43
	D2	-286.24*	55.583	.000	-453.25	-119.23
	D1	-78.03	59.808	.848	-257.73	101.67
	L	-33.90	65.289	.999	-230.07	162.27
	A1	-61.33	66.560	.968	-261.32	138.66
	A2	-47.52	61.318	.987	-231.76	136.72

Based on observed means. The error term is Mean Square(Error) = 40221.934.

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

Table F4 Tukey HSD table of mean accuracy of return to task one in ms

Behavior (I)	Behavior (J)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
D3	D2	452.29	398.46	0.917	-726.86	1631.44
	D1	320.34	389.76	0.983	-833.06	1473.75
	L	-508.62	386.41	0.844	-1652.10	634.86
	A1	-1549.19	388.62	0.001	-2699.22	-399.16
	A2	-1435.12	385.33	0.004	-2575.41	-294.82
	A3	-2450.55	381.24	0.000	-3578.74	-1322.37
D2	D3	-452.29	398.46	0.917	-1631.44	726.86
	D1	-131.94	399.57	1.000	-1314.39	1050.50
	L	-960.91	396.30	0.190	-2133.67	211.85
	A1	-2001.48	398.46	0.000	-3180.63	-822.33
	A2	-1887.40	395.25	0.000	-3057.06	-717.74
	A3	-2902.84	391.26	0.000	-4060.69	-1744.99
D1	D3	-320.34	389.76	0.983	-1473.75	833.06
	D2	131.94	399.57	1.000	-1050.50	1314.39
	L	-828.97	387.55	0.331	-1975.84	317.91
	A1	-1869.53	389.76	0.000	-3022.94	-716.13
	A2	-1755.46	386.48	0.000	-2899.17	-611.75
	A3	-2770.90	382.40	0.000	-3902.52	-1639.27
L	D3	508.62	386.41	0.844	-634.86	1652.10
	D2	960.91	396.30	0.190	-211.85	2133.67
	D1	828.97	387.55	0.331	-317.91	1975.84
	A1	-1040.57	386.41	0.102	-2184.05	102.91
	A2	-926.49	383.10	0.193	-2060.18	207.20
	A3	-1941.93	378.98	0.000	-3063.44	-820.43
A1	D3	1549.19	388.62	0.001	399.16	2699.22
	D2	2001.48	398.46	0.000	822.33	3180.63
	D1	1869.53	389.76	0.000	716.13	3022.94
	L	1040.57	386.41	0.102	-102.91	2184.05
	A2	114.07	385.33	1.000	-1026.22	1254.37
	A3	-901.36	381.24	0.216	-2029.55	226.82
A2	D3	1435.12	385.33	0.004	294.82	2575.41
	D2	1887.40	395.25	0.000	717.74	3057.06
	D1	1755.46	386.48	0.000	611.75	2899.17
	L	926.49	383.10	0.193	-207.20	2060.18
	A1	-114.07	385.33	1.000	-1254.37	1026.22
	A3	-1015.44	377.88	0.103	-2133.70	102.82
A3	D3	2450.55	381.24	0.000	1322.37	3578.74
	D2	2902.84	391.26	0.000	1744.99	4060.69
	D1	2770.90	382.40	0.000	1639.27	3902.52
	L	1941.93	378.98	0.000	820.43	3063.44
	A1	901.36	381.24	0.216	-226.82	2029.55
	A2	1015.44	377.88	0.103	-102.82	2133.70

Based on observed means. The error term is Mean Square(Error) = 6494088.904.

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

Table F5 Tukey HSD table of number of mean successful completions of task two per trial

Behavior (I)	Behavior (J)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
D3	D2	.57	.349	.660	-.46	1.60
	D1	-.49	.349	.804	-1.52	.54
	L	-1.68*	.349	.000	-2.71	-.65
	A1	-2.63*	.349	.000	-3.66	-1.60
	A2	-2.46*	.349	.000	-3.50	-1.43
	A3	-3.06*	.349	.000	-4.10	-2.03
D2	D3	-.57	.349	.660	-1.60	.46
	D1	-1.06*	.349	.040	-2.09	-.03
	L	-2.25*	.349	.000	-3.29	-1.22
	A1	-3.2*	.349	.000	-4.23	-2.17
	A2	-3.04*	.349	.000	-4.07	-2.00
	A3	-3.63*	.349	.000	-4.67	-2.60
D1	D3	.49	.349	.804	-.54	1.52
	D2	1.06*	.349	.040	.03	2.09
	L	-1.2*	.349	.012	-2.23	-.16
	A1	-2.14*	.349	.000	-3.18	-1.11
	A2	-1.98*	.349	.000	-3.01	-.95
	A3	-2.58*	.349	.000	-3.61	-1.55
L	D3	1.68*	.349	.000	.65	2.71
	D2	2.25*	.349	.000	1.22	3.29
	D1	1.2*	.349	.012	.16	2.23
	A1	-.95	.349	.095	-1.98	.08
	A2	-.78	.349	.275	-1.81	.25
	A3	-1.38*	.349	.002	-2.41	-.35
A1	D3	2.63*	.349	.000	1.60	3.66
	D2	3.2*	.349	.000	2.17	4.23
	D1	2.14*	.349	.000	1.11	3.18
	L	.95	.349	.095	-.08	1.98
	A2	.17	.349	.999	-.87	1.20
	A3	-.43	.349	.878	-1.46	.60
A2	D3	2.46*	.349	.000	1.43	3.50
	D2	3.04*	.349	.000	2.00	4.07
	D1	1.98*	.349	.000	.95	3.01
	L	.78	.349	.275	-.25	1.81
	A1	-.17	.349	.999	-1.20	.87
	A3	-.60	.349	.604	-1.63	.43
A3	D3	3.06*	.349	.000	2.03	4.10
	D2	3.63*	.349	.000	2.60	4.67
	D1	2.58*	.349	.000	1.55	3.61
	L	1.38*	.349	.002	.35	2.41
	A1	.43	.349	.878	-.60	1.46
	A2	.60	.349	.604	-.43	1.63

Based on observed means. The error term is Mean Square(Error) = 6.331.

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

Table F6 Tukey HSD table of log transform of mean length of time before first switch to task two in ms

Progress Bar Behavior (I)	Progress Bar Behavior (J)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Decelerating	Linear	.0715*	0.016	0.000	0.034	0.109
	Accelerating	0.032	0.015	0.098	-0.004	0.068
Linear	Decelerating	-.0715*	0.016	0.000	-0.109	-0.034
	Accelerating	-.0396*	0.016	0.033	-0.077	-0.003
Accelerating	Decelerating	-0.032	0.015	0.098	-0.068	0.004
	Linear	.0396*	0.016	0.033	0.003	0.077

Based on observed means. The error term is Mean Square(Error) = .011.

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

Table F7 Tukey HSD table of mean number of checks on the progress bar per trial

Progress Bar Behavior (I)	Progress Bar Behavior (J)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Decelerating	Linear	.3226*	0.050	0.000	0.205	0.441
	Accelerating	.3226*	0.050	0.000	0.205	0.441
Linear	Decelerating	-.3226*	0.050	0.000	-0.441	-0.205
	Accelerating	0.000	0.050	1.000	-0.118	0.118
Accelerating	Decelerating	-.3226*	0.050	0.000	-0.441	-0.205
	Linear	0.000	0.050	1.000	-0.118	0.118

Based on observed means. The error term is Mean Square(Error) = .130.

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

Table F8 Tukey HSD table of mean length of check on progress bar in ms

Progress Bar Behavior (I)	Progress Bar Behavior (J)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Decelerating	Linear	225.9322*	83.303	0.021	27.554	424.310
	Accelerating	18.135	62.857	0.955	-131.553	167.823
Linear	Decelerating	-225.9322*	83.303	0.021	-424.310	-27.554
	Accelerating	-207.7968*	87.097	0.049	-415.211	-0.383
Accelerating	Decelerating	-18.135	62.857	0.955	-167.823	131.553
	Linear	207.7968*	87.097	0.049	0.383	415.211

Based on observed means. The error term is Mean Square(Error) = 105743.441.

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

Table F9 Tukey HSD table of mean accuracy of return to task one in ms

Progress Bar Behavior (I)	Progress Bar Behavior (J)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Decelerating	Linear	-736.849	348.800	0.089	-1559.355	85.658
	Accelerating	-2392.6071*	341.254	0.000	-3197.321	-1587.893
Linear	Decelerating	736.849	348.800	0.089	-85.658	1559.355
	Accelerating	-1655.7585*	347.934	0.000	-2476.223	-835.294
Accelerating	Decelerating	2392.6071*	341.254	0.000	1587.893	3197.321
	Linear	1655.7585*	347.934	0.000	835.294	2476.223

Based on observed means. The error term is Mean Square(Error) = 5618775.457.
*. The mean difference is significant at the .05 level.

Table F10 Tukey HSD table of number of mean successful completions of task two per trial

Progress Bar Behavior (I)	Progress Bar Behavior (J)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Decelerating	Linear	-1.8670*	0.296	0.000	-2.564	-1.170
	Accelerating	-2.8969*	0.296	0.000	-3.594	-2.199
Linear	Decelerating	1.8670*	0.296	0.000	1.170	2.564
	Accelerating	-1.0299*	0.296	0.002	-1.727	-0.332
Accelerating	Decelerating	2.8969*	0.296	0.000	2.199	3.594
	Linear	1.0299*	0.296	0.002	0.332	1.727

Based on observed means. The error term is Mean Square(Error) = 4.555.
*. The mean difference is significant at the .05 level.

Appendix G

INFORMED CONSENT FORM FOR BEHAVIORAL RESEARCH STUDY
Rochester Institute of Technology

Title of Project:	**Time Delays and System Response Times in Human-Computer Interaction**	
Investigators in Charge:	Mr. Noah Stupak	Dr. Esa M. Rantanen
	MS Candidate	Associate Professor
	Dept. of Psychology.	Dept. of Psychology
	Rochester Inst. of Technology	Rochester Inst. of Technology
	Tel. (585)414-1966	01-3140 Eastman Bldg.
	Email: njs4257@rit.edu	Tel. (585) 475-4412
		Email: esa.rantanen@rit.edu

Explanation of the Project.

1. You are being asked to participate in a research study that is looking at the representation of time on a computer. The results of this study will be applicable to HCI where the representation of time is necessary. Design guidelines will be developed that may improve user performance and satisfaction, with the greatest benefit to medical and flight computers, where representation of time is of greatest consequence

2. The goal of this work is to evaluate humans' ability to asses system response times presented graphically.

3. This study requires you to engage in two simultaneous tasks, which take the form of windows, viewable one at a time. You will be responsible for (1) a search task, which entails searching a database for items, and (2) a visual search task, which will require you to search for a target or indicate that the target is not present. Timely completion of the search task is imperative.

4. The only risks to you from participating in the experiment are the slight mental workload and fatigue associated with any search task.

5. Results of this research will be used to further enhance our understanding of the role of time in human performance.

Your rights as a research participant

1. We will be happy to answer any questions you have about the study at any time. Mr. Stupak and Prof. Rantanen may be contacted at the telephone numbers and e-mail addresses shown above. If you have questions about your rights as a research subject, you can call collect the Rochester Institute of Technology Institutional Review Board at (585) 475-7673, or e-mail hmfsrs@rit.edu.

2. No subsequently published results will contain any information that could be associated with individual participants. No information identifying individual subjects will be ever associated with the data collected. All data will be stored and secured only on the investigator's computer after being retrieved from the program.

3. Your participation is wholly voluntary. Your decision to participate, or to not participate, or to withdraw from the study during the experiment will in no way influence your relationship with the researcher or your professor(s).

4. You may refuse to participate or may discontinue participation at any time during the project without penalty or loss of benefits to which you are otherwise entitled.

5. Results of the proposed research will be used to further guide our understanding of temporal awareness.

6. The results of this research will be submitted to peer-reviewed journal articles and perhaps presented at a human factors-related conference. No information allowing for identification of individual participants will be included in these reports.

Statement of consent

Participant:

I agree to participate in this study, which seeks to guide development and testing of human performance in supervisory, time-sensitive environs. I understand the information given to me, and I have received answers to any questions I may have had about the research procedure. I understand and agree to the conditions of this study as described on this form.

I understand that I am volunteering to participate in this study, that I will be not be compensated for participating apart from the chances of winning a raffle, and that I may withdraw from this study at any time without penalty to me.

I certify that I am at least 18 years old.

I understand that I will be given a signed copy of this consent form.

____________________________________ _______________________

Signature Date

Researcher:

I certify that the informed consent procedure has been followed, and that I have answered any questions from the participant above as fully as possible.

____________________________________ _______________________

Signature Date

Appendix H

R·I·T

Rochester Institute of Technology

RIT Institutional Review Board for the
Protection of Human Subjects in Research
141 Lomb Memorial Drive
Rochester, New York 14623-5604
Phone: 585-475-7673
Fax: 585-475-7990
Email: hmfsrs@rit.edu

Form C
IRB Decision Form

TO: Noah Stupak; Esa Rantanen

FROM: RIT Institutional Review Board

DATE: **9/24/08**

RE: **Decision of the RIT Institutional Review Board**

Project Title – The Effect of Displays of System Response Time on User Performance in Human-Computer Interaction

The Institutional Review Board (IRB) has taken the following action on your project named above.

☒ Approved, no greater than minimal risk

Now that your project is approved, you may proceed as you described in the Form A. **Note that this approval is only for a maximum of 12 months; you may conduct research on human subjects only between the date of this letter and 9/24/09.**

You are required to submit to the IRB any:

- **Proposed** modifications and wait for approval before implementing them,
- Unanticipated risks, and
- Actual injury to human subjects.

Return the Form F, at the end of your human research project or 12 months from the above date. If your project will extend more than 12 months, your project must receive continuing review by the IRB.

Heather Foti

Heather Foti
Associate Director, Office of Human Subjects Research

Appendix I

Figure I1 Boxplot of log transform of length of time before switching to task two per condition.

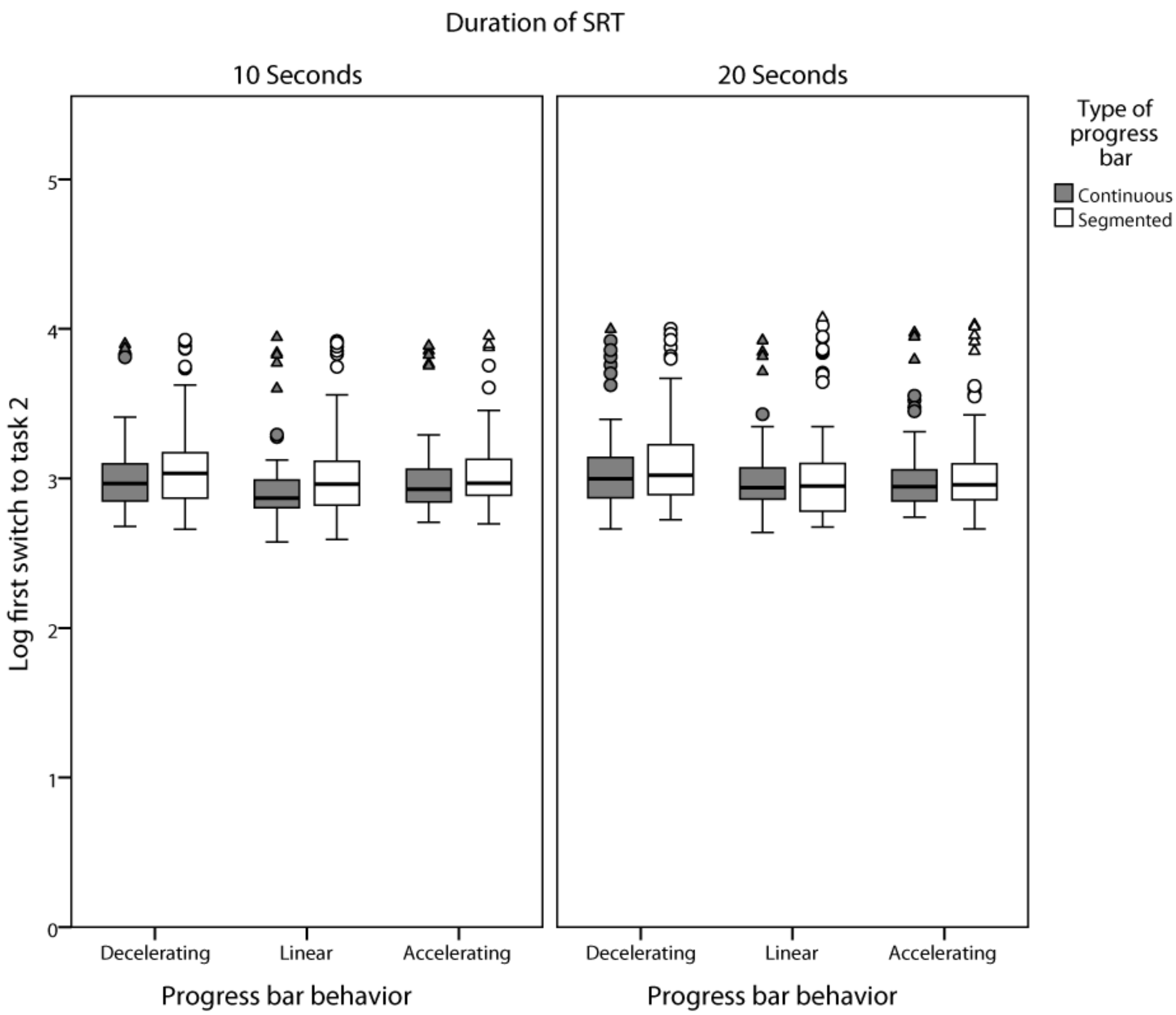


Figure I2 Boxplot of number of successful completions of task two per condition.

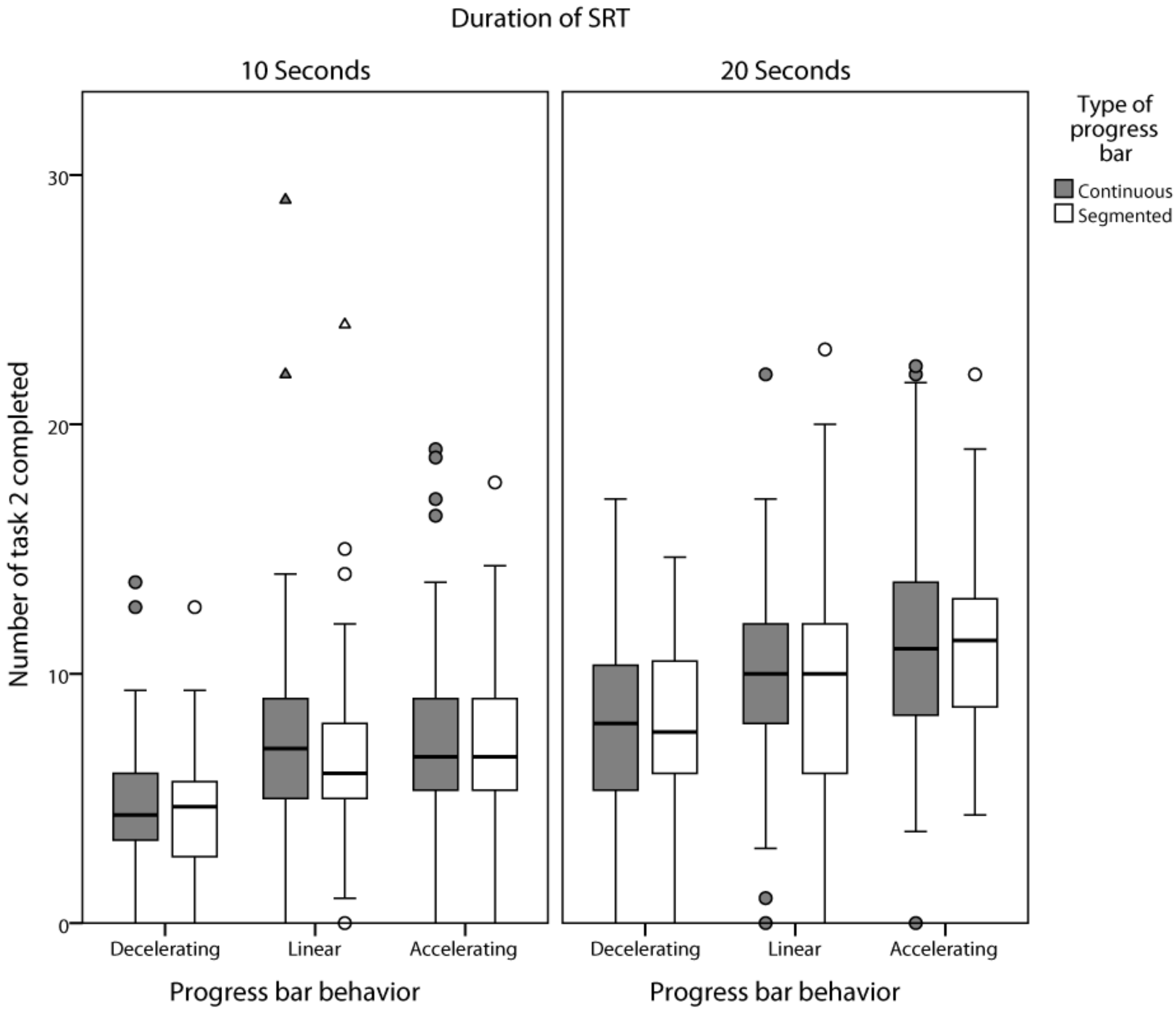


Figure I3 Boxplot of number of checks on the progress bar per condition.

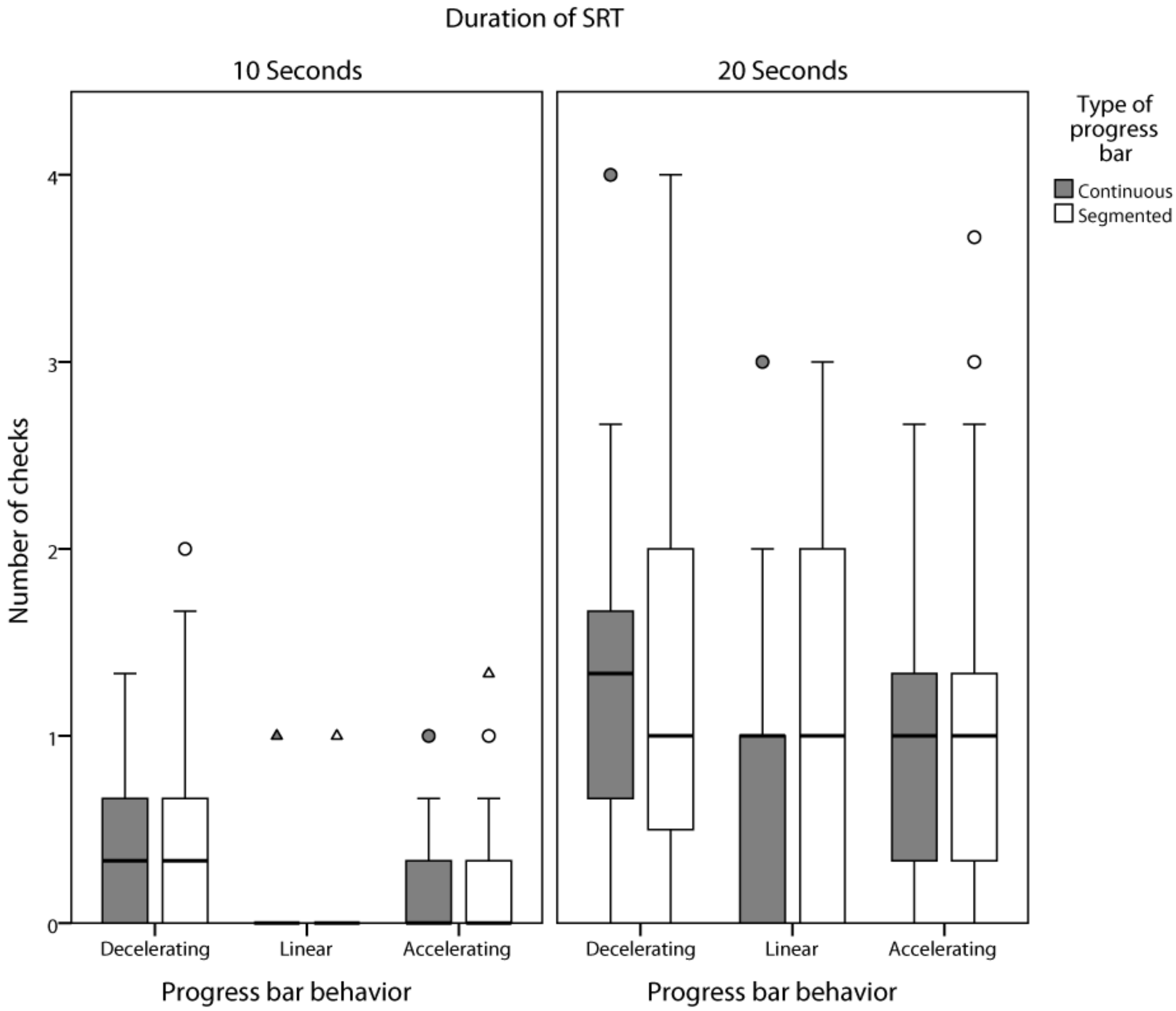


Figure I4 Boxplot of length of check on the progress bar per condition.

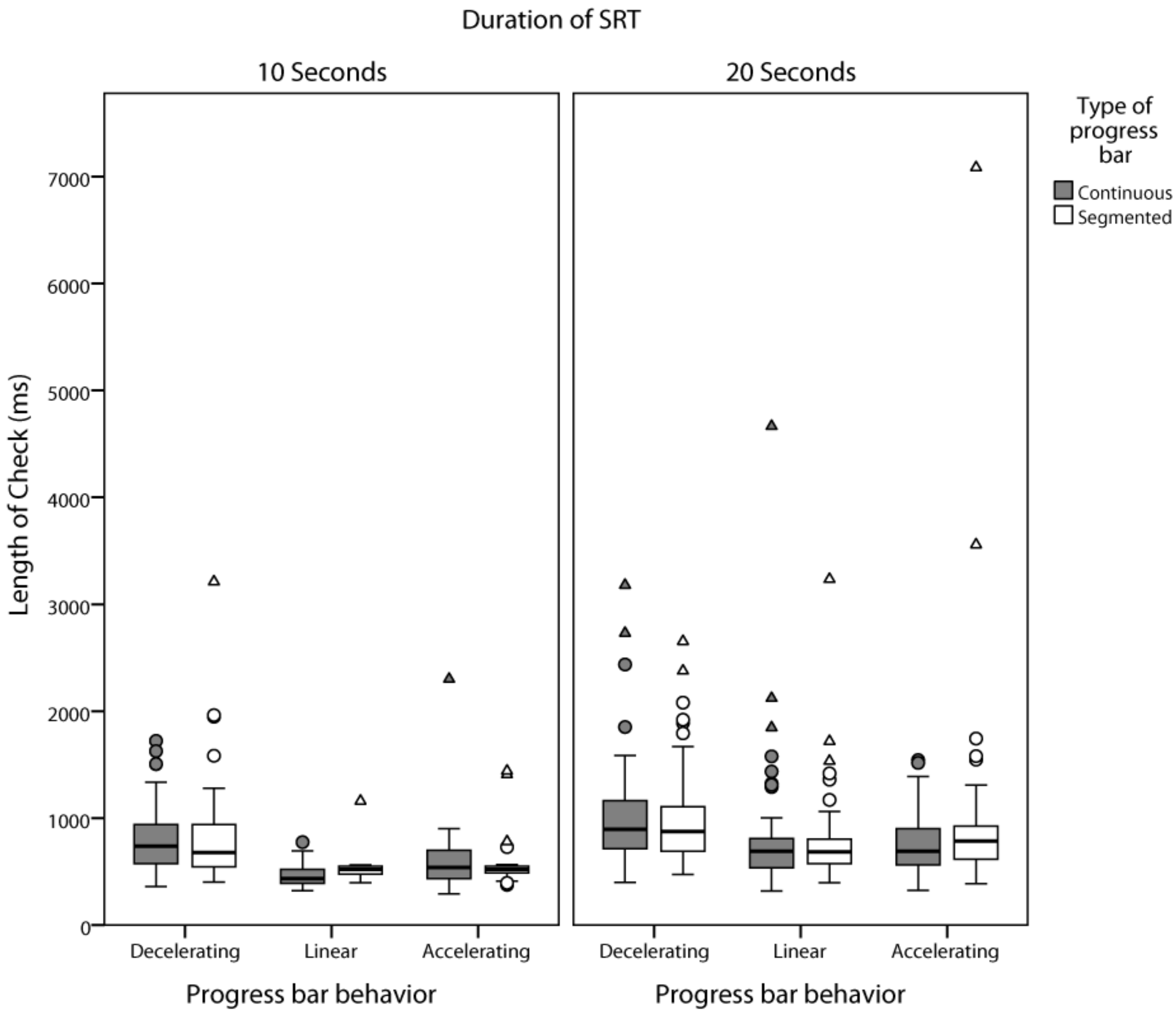


Figure I5 Boxplot of error of return to primary task per condition.