AFIT-ENV-DS-20-M-205

EVALUATION OF TEXT-BASED INFORMATION PORTRAYAL AND ACCESS TECHNIQUES FOR CONSTRAINED AREA PRESENTATION: A COMPARISON OF HUMAN PERFORMANCE

DISSERTATION

Eric E. Geiselman, Civilian, USAF

AFIT-ENV-DS-20-M-205

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENV-DS-20-M-205

AFIT-ENV-DS-20-M-205

EVALUATION OF TEXT-BASED INFORMATION PORTRAYAL AND ACCESS TECHNIQUES FOR CONSTRAINED AREA PRESENTATION: A COMPARISON OF HUMAN PERFORMANCE

DISSERTATION

Presented to the Faculty

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

Eric E. Geiselman, BS, MA

Civilian, USAF

March 2020

AFIT-ENV-DS-20-M-205

EVALUATION OF TEXT-BASED INFORMATION PORTRAYAL AND ACCESS TECHNIQUES FOR CONSTRAINED AREA PRESENTATION: A COMPARISON OF HUMAN PERFORMANCE

Eric E. Geiselman, BS, MA

Civilian, USAF

Committee Membership

Dr. Michael E. Miller
Chair

Dr. John M. Colombi
Member

Dr. Christine M. Schubert Kabban
Member

Dr. Adedeji B. Badiru, PE
Dean, Graduate School of Engineering and Management

Abstract

To maximize visualization effectiveness, graphical data are commonly augmented with alphanumeric (i.e., text) symbols to provide detailed information and define specific values. This text is often provided through a pop-up dialog box which contains all of the alphanumeric data pertaining to an object simultaneously. However, a human performance cost can be associated with the current portrayal and access techniques as the resulting portrayal can occlude critical information in the visualization and degrade operator performance. This research included the development and evaluation of three alternate spatially-constrained text portrayal techniques. These techniques and their associated access interface were designed to reduce occlusion while providing rapid access to desired alphanumeric data. The techniques were comparatively evaluated using a dual-task human performance paradigm. Among the performance measures were accuracy, response time, and subjective feedback. The basis for development of the spatially-constrained text access techniques is discussed as well as their implementation affordances and limitations. The spatially-constrained text portrayal and access user interface concepts resulted in mixed accuracy and response time performance compared with the more conventional method. Specific design features promoted a 3X reduction in data access time versus unaided spatially-constrained text portrayal. Overall, it was shown that equivalent performance was obtained among the variants while the potential for occlusion was reduced during use of the novel designs.

Acknowledgments

I express my sincere gratitude to all of the members of my committee for their valuable input and guidance throughout this process. I very much appreciate their time and effort expended on my behalf. To Dr. Miller, thank you for never complaining to me about the number of drafts I sent your way. And thank you also for always sending an improvement back to me. I thank Mr. Eric Heft of the Air Force Research Laboratory's 711th Human Performance Wing for his expert coding of these experiments and associated data collection routines. Thank you to Mr. Randy Brown for getting me into this… I'll get you back. Thank you to Mr. Jeff Craig, Dr. Paul Havig, and Dr. Fairul Mohd-Zaid for all the encouragement and help over the years. And to Dr. George Reis, you're a splendid partner in crime. My appreciation to Ms. Lauren Menke for her assistance during subject scheduling and data collection. Also, I very much appreciate the time and effort of all those who volunteered to participate in these studies. Thank you to my wife Kathryn for never asking me when I would finish this. Thank you to my father-in-law Mr. Perry Bradley Jr. for always wanting to know when I would finish it. Thank you to my mother Judith for never letting us realize how hard she worked for us. To my daughters Ellen and Lily, thanks for making me forever proud.

Eric E. Geiselman

Table of Contents

Page

List of Figures

List of Tables

Page

List of Abbreviations

ANOVA	………………..	Analysis of Variance
AWACS	………………..	Airborne Warning and Control System
BP	………………..	Battle Point
C^3I	………………..	Control, Communication, and Intelligence
CAS	………………..	Close Air Support
CC	………………..	Cinema Credits
cm	………………..	Centimeter
CRT	………………..	Cathode Ray Tube
d	………………..	Difficulty
DB	………………..	Dialog Box and Pop-Up Dialog Box
DD	………………..	Dynamic Dependence
deg	………………..	Degrees
DoD	………………..	Department of Defense
er	………………..	Event Rate
FOV	………………..	Field of View
GRE	………………..	Graduate Record Exam
h	………………..	Horizontal
HMD	………………..	Helmet-Mounted Display
HSD	………………..	Honest Significant Difference
ICD	………………..	Informed Consent Document
ID	………………..	Identification

IES	………………..	Inverse Efficiency Score
IP	………………..	Initial Point
IRB	………………..	Institutional Review Board
ISI	………………..	Inter-Stimulus Interval
LAD	………………..	Large Area Display
lat	………………..	Latitude
long	………………..	Longitude
M	………………..	Mean
MAT-B	………………..	Multi-Attribute Task Battery
ms	………………..	Millisecond
MSL	………………..	Mean Sea Level
n	………………..	Number
p	………………..	Probability
PDA(s)	………………..	Personal Digital Assistant(s)
PID	………………..	Presence, Identification, Direction
r_s	………………..	Spearman's Rank Correlation Coefficent
RAP-COM	………………..	Rapid Communication Display
RSVP	………………..	Rapid Serial Visual Processing
s	………………..	Second
SC	………………..	Spatially-Constrained
SI	………………..	Static Independence
SOA	………………..	Stimulus Onset Asynchrony

STM	………………..	Short-Term Memory
SysML	………………..	Systems Modeling Language
TRL(s)	………………..	Technology Readiness Level(s)
TS	………………..	Times Square
UTC	………………..	Universal Time Constant
v	………………..	Vertical
WPM	………………..	Words-Per-Minute

EVALUATION OF TEXT-BASED INFORMATION PORTRAYAL AND ACCESS TECHNIQUES FOR CONSTRAINED AREA PRESENTATION: A COMPARISON OF HUMAN PERFORMANCE

I. Introduction

General Issue

The fundamental objective of information visualization is to afford human operators an appreciation of complex relationships among data in a way that is accessible and easy to interpret. The overall goal is to support interpretation accuracy and enable timely understanding in terms of meaning making. Within the general utility of visualization, a high-level distinction can be made between portrayals intended to fulfill the objectives of information discovery (e.g., scatter plots of statistical relationships) or information interpretation (e.g., representations of real-world events and timelines). Certain visualization formats lend themselves more toward one objective versus the other. Also, visualization format choice and its associated information extraction value is similarly task dependent. From the operator or user perspective, visualization format usefulness depends on how well the information presented supports the process of effective (accurate and timely) decision making. Information accessibility, accuracy, interactivity, and timeliness are also desirable across a wide variety of portrayal approaches. These characteristics in proper combination form the overall effectiveness and efficiency of a visualization design and are strongly related to human operator performance. To realize effective visualization, it is not uncommon that graphical data visualization presentation be augmented with alphanumeric (text) symbols for the purpose of labeling, detail presentation, defining specific values, etc.

This research applies to electronic display media. More specifically, both dynamic display and interactivity between the intended user and the data are assumed. A good metaphorical example of the target visualization of interest is that used in Command, Control, Communication, and Intelligence (C^3I) applications. Examples of these types of displays are: air traffic control large-area radar depictions, Airborne Warning and Control System (AWACS) battlespace visualization workstation displays, cyber activity monitoring graphs, and large-area situation displays. Figure 1 shows the type of operator workstation under consideration. Figure 2 is an example of the type of complexity which is often found in existing visualizations. Within the given example in Figure 3, a "pop-up dialog box" is used to provide text-based information regarding a selected entity of interest. In that example, the pop-up dialog box is shown occluding a portion of the underlying display.

Figure 1. Command and control operator workstation example (U.S. Air Force, 2011).

Looking toward future applications, visualization requirements in support of multiple remotely piloted aircraft supervisory control, space situation awareness, and flightdeck tactical situation awareness displays all pose similar challenges. For the user, as a decision maker, the needs are common among these examples. The display should support periodic shifts of user attention away from the wide area “big picture” appreciation of dynamic spatio-temporal relationships among entities of interest, traditionally supported by a “God’s eye view” of the scenario environment. The display should also support the ability to “drill down” to obtain detailed information pertaining to entities of interest. Of course the continuum between these extremes must be accessible as well. Simultaneously, any new presentation format must still adhere to the principles of readability that have evolved over time. Moreover, to be able to determine if novel formats provide advantage over current approaches, empirical validation methods are required during usability evaluations (Öquist, 2006).

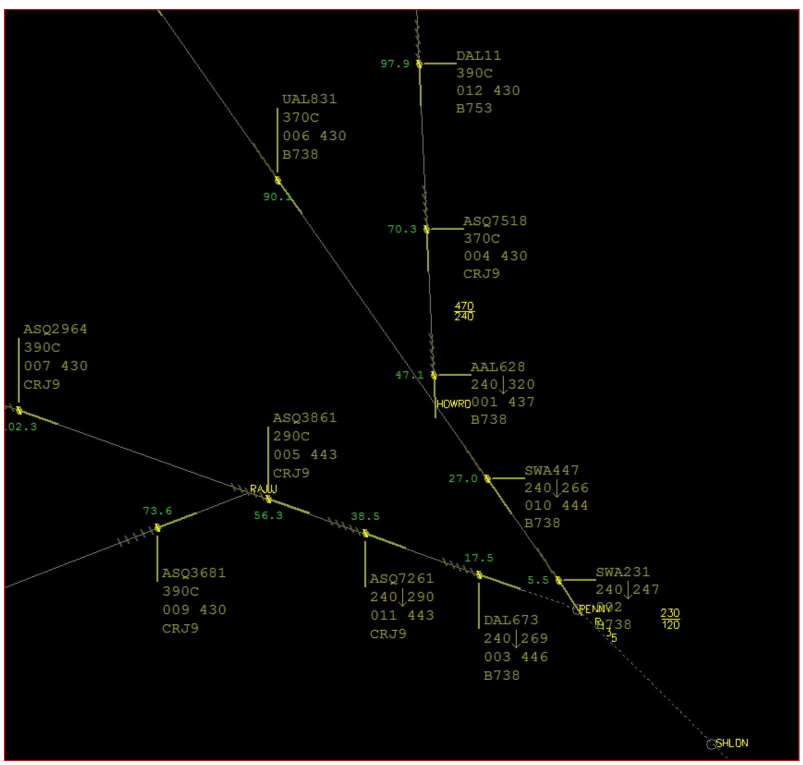


Figure 2. Complex data display and overlapping text (Federal Aviation Administration).

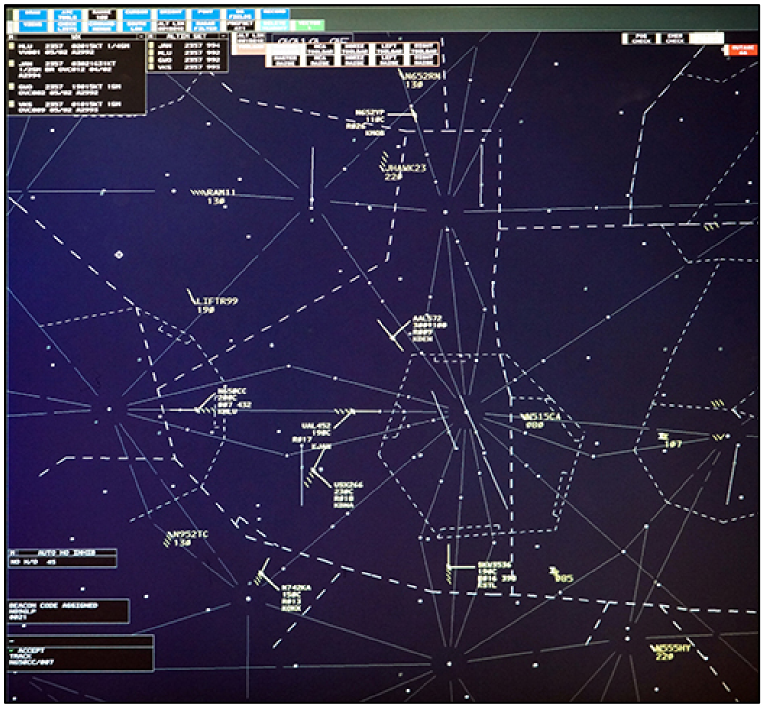


Figure 3. Complex visualization with pop-up dialog box text (Federal Aviation Administration, 2018).

Problem Statement

The desire to simultaneously provide “big picture” visualization with the ability to access detailed text-based information is ubiquitous among visualization applications. However, the human performance cost associated with current portrayal and access techniques may be unacceptable if these approaches preclude the operator from maintaining the “big picture”. The current research seeks to demonstrate methods which reduce such human performance cost via a systematic design and evaluation process that employs a human-systems integration technical development approach.

General Research Hypotheses

1) Various information portrayal techniques can be devised which will result in measurable human performance differences during otherwise identical text access tasks.

2) Portrayal techniques which present all information simultaneously will result in accurate and quick access to text information in comparison to...

3) Portrayal techniques which occlude information will interfere with and produce measurable degradation of tasks that depend on the use of the occluded information.

4) Optimal information presentation is portrayal feature (technique) dependent.

Specified Research Hypotheses

1) Novel information portrayal and interaction techniques can be designed which will result in enhanced human performance during otherwise identical text access tasks when compared to a technique representative of current capability (baseline technique).

1a) Given the application of a secondary task paradigm methodology and the associated prioritization of the primary task, there will be no statistically reliable error or response time differences among the manipulation of text access techniques within the primary task.

1b) Given the application of a secondary task paradigm methodology and the associated prioritization of the primary task, there will be a statistically reliable error reduction as a result of novel text access technique use when compared to performance using the baseline technique measured via the secondary task.

1c) Given the prioritization of task accuracy versus response time, response time will not differ statistically significantly across any comparison for correct responses.

1d) As a result the user interaction requirements of the novel text access techniques, reported subjective workload will be statistically significantly higher compared to the baseline technique but will remain acceptable.

2) The baseline technique which presents all information simultaneously will result in accurate and quick access to the text information but will present an occlusion problem.

2a) The increased display area required of the baseline text access technique will result in a statistically reliable decrease in secondary task accuracy due to the occlusion of events of interest compared to the novel designs.

2b) Due to the prioritization of accuracy for both the primary and secondary tasks, response time for correct responses will not provide statistically reliable differences across text access technique.

3) A portrayal technique which occludes information will interfere more with tasks that depend on the use of the occluded information compared to novel access techniques that have been designed to minimize occlusion.

3a) The measurement of occlusion during performance of the experimental tasks will result in statically significant higher potential for secondary task error during use of the baseline text access technique compared to use of the novel designs.

4) A novel design feature which affords text access to remain physically closer to dynamic events of interest will generate advantageous performance compared to that which allows the separation of the events of interest from the text access task.

    4a) Text access attached to the dynamic entities will result in statistically significantly fewer secondary task errors when compared to text access fixed in screen coordinates.

    4b) The manipulation of text access separation will not result in a statistically significant difference in reported subjective workload.

Research Focus

In terms of effectiveness, the need to incorporate text into a complex visualization may have unintended consequences. This is especially true considering the relatively large display surface area or "real estate" required to ensure that visually displayed text be readable via the unaided human visual system. Either presented on the written page or through electronic display media, minimum levels of brightness, contrast, character size, and spacing must be achieved for readability to be acceptable (Kruk & Muter, 1984). The value of visualization can be summed up with the classic axiom: "a picture is worth a thousand words," but what must a word be worth to justify including it with a picture?

In some applications, there is often more information the interface designer wishes to make accessible to the user than there is physical display surface area available to avoid unacceptable occlusion. Regardless of the display scale under consideration, be it a data wall or a mobile telephone screen, when text presentation is deemed necessary, the effectiveness with which it can be included in the visualization is an important design challenge. If a large amount of text is displayed at once, such as inside a pop-up dialog box, significant occlusion cost may be incurred. When attempting to display text within a small area, a readability and/or accessibility cost may be incurred.

The objective of this research is to design and evaluate generalizable competing methods to afford operator access to text-based information within operationally representative complex and dynamic visualizations. The concepts of interest share a simple design question: how can the amount of readable text-based information available to a user be maximized while minimizing the display surface area required for its availability and accessibility? Can human performance, specifically visualization effectiveness, be enhanced by designing a means for effective access to text while utilizing a minimal amount of display real estate?

Considering the physiology of the human visual system, it is the case that only a relatively small amount of displayed detail can be resolved during a single fixation. As presented by Cornsweet (1970), the cone type photoreceptors of the human retina are exclusive to and densely packed into an area called the fovea. Because of the density of the cones within the foveal area (approximately 140,000 cones per mm^2 within one degree of redial area at the center of retina), and the optical physics of the human eye which focuses incoming light onto the fovea, the fovea provides the physiological basis for human perception input of fine detailed visual information. Regardless of the physical display surface size, at any given distance from the display and under adequate luminance levels, the human visual system is able to resolve the highest level of information detail at approximately just less than 4 degrees subtended visual angle. Detail falls off significantly and rapidly across the next 5 degrees (Grand, 1975). To “see” the detail information content of a display area such as the examples presented above, the area must be scanned by a series of eye movements with associated sampling fixation points. The eye movement between the fixation points is called a saccade. The perceptual phenomena described here have implications regarding how humans read text as well as how the overall content of displayed information is processed visually.

The focus of the present research is an empirical investigation of the relative merits of different techniques operators may use to view and access text-based information in the context of operationally representative tasks involving the display of complex visual information. The research includes the conceptual development of the text access techniques as well as the operationally representative evaluation environment and task scenario. Human performance measurements, typically task accuracy and completion duration (response time), are the dependent variables of interest. This effort includes the balance of experimental control for purposes of isolating any performance effects attributable to the independent variables of interest with the desire to generalize the findings to other applied operational environments. An objective of the research is the demonstrated advancement of the technology (“what works?” and “what doesn’t work?”) enhanced by the advancement of general knowledge of the application for what works and what doesn’t (“why did performance differ?”).

Investigative Questions

1) What is the best way to present the most amount of usable text-based information within a constrained amount of display real estate?

2) How best should operators be given access to the content of text-based information when the spatial dimension for its presentation is constrained?

3) What are the relative advantages and disadvantages of the various mechanizations within an operationally representative task environment?

4) How well does the methodology measure the human performance differences resulting from the variations among the text access designs and information portrayal features?

Methodology Preview

This research employed a dual- (also called secondary-) task paradigm as the basis of an empirical evaluation (Gawron, 2008). The primary task was designed to exercise the utility of the text access independent variables of interest. Participants used different text access techniques to respond to queries where the task required locating and verifying information content such as specific data values. The secondary task was designed to be a sensitive performance measurement activity (accuracy and response time) made up of monitoring moving entities for state changes and then reporting those changes. Secondary task difficulty was determined empirically via Experiment 1 presented in Chapter III of this document. Both the primary and secondary tasks competed for the same perceptual-motor resources. Based on the sensitivity of the secondary task, qualitative differences among the text access technique affordances should systematically interfere with participants' ability to perform the secondary task such that better secondary task performance was intended to be indicative of a "better" text access technique interface (i.e., better human-system integration). Thus, this enabled the differences to be measured objectively and the differences to be analyzed using descriptive and inferential statistical methods.

Subjective measures in the form of workload assessment and preference questionnaires were also administered to help interpret the performance measures described above.

The overall secondary task methodology and outcome (Experiment 2) is presented in Chapter IV of this document.

Assumptions and Limitations

The body of text utilized for the experiments was formatted in a standardized sequence. This was both an assumption as well as a limitation of the usefulness of the design candidates. Fortunately, standardized presentation is often the case for operational systems. For instance, data blocks for aircraft returns within air traffic control digital radar systems are presented in a standardized form. Figure 2 and Figure 3 are good examples of this where the number of lines of presented text is standardized for the displayed entities.

Another limitation is, for reasons of experimental control and to afford the use of individuals who are not subject matter experts as study participants (trained naïve participants), an operationally "representative" task was developed as the basis for this research. In terms of generalizing the findings, the design concepts will need to be further demonstrated at advanced Technology Readiness Levels (TRLs). Of course this is the normal progression of technology development and transition but the present research was most appropriately assigned a TRL of late 3 to early 4 (DOD, 2010). That is, the research was intended as an analytical proof-of-concept in a simulated environment.

Implications

Real-world visualization applications invariably demonstrate that text-based information is necessary and present in some form. The text-based information affords the user detailed and specific data about entities of interest within the larger scenario. The successful development and evaluation of text access techniques which were designed to minimize interference with the "big-picture" appreciation of the scenario should promote enhanced overall performance for both levels of information (global "big picture" and text based detail on demand) and thus, promote more effective decision making.

Preview

The following literature review is intended to provide the basis and motivation for this research and provide the conceptual basis for the candidate human system integration techniques evaluated here. The experimental methodology utilized is then presented including descriptions of the text access techniques. The evaluation approach is illustrated through a preliminary study which was conducted to determine the appropriate level of difficulty and workload for the secondary task component. The full study methodology is presented as well as the objective and subjective results. The document concludes with an interpretation of the findings in terms of application implications, limitations, design recommendations, and suggestions for further study.

II. Literature Review

Chapter Overview and Objective

The following chapter presents existing research which was influential on the methodology utilized or included previously identified design and information portrayal principles. These later principles were incorporated into the competing designs that were developed and evaluated within this dissertation.

Reading and Performance Measurement

The act of reading text is typically a visual task. More specifically, the mechanisms involved include the foveal region of the human visual system for symbol detection and differentiation, and the language areas of the brain for the cognitive processing required for interpretation and meaning making. From a bottom-up perspective, the visual input of the reading process can be described in psychophysical terms. For symbols, and ultimately words, to be perceived as input for the information processing loop, proper conditions facilitating disambiguation must be satisfied. Furthermore, a scanning and sampling behavior is required for this input to be successful.

For a conventional reading task, such as reading the text on this page, the reader is required to make periodic saccadic eye movements between fixations. The saccadic eye movements are swift, ballistic eye movements made to move the fovea across the page in a fairly regular, repetitive pattern to land the foveal region on a series of sequential targets (Rayner, 1998). Fixations are short pauses of the eye which permit the collection of detailed visual information. As the interpretation of text requires the differentiation of high spatial frequency information, reading requires the high resolution of the fovea to differentiate symbols. Therefore, saccadic eye movements occur while reading a line of text which are typically separated by about 7-9 letter spaces, occurring about every 2 degrees subtended visual angle for an approximately 12 point font at a normal reading distance (Morrison & Rayner, 1981; Rayner, 1998).

Processing from the top-down includes symbol recognition, pattern interpretation, memory-matching (knowledge, expertise, vocabulary, etc.), and ultimately meaning assignment toward the completion of a language acquisition task. This is the process of reading related to the “context” of the subject matter (Lindsay & Norman, 1972; Wickens, Banbury, & Parasuraman, 2013). When combined, these processes form the basis of reading comprehension which, when measured, is used as quantitative metric of observer understanding. This is often expressed as a percent of correctly answered questions, where the questions are designed to probe the reader’s understanding of the text (Öquist, 2006) or the reader’s ability to correctly identify a word or phrase that is grammatically appropriate but nonsensical within the structure of a paragraph (Chapman & Cook, 1923).

Reading efficiency as a measurement is typically temporally based, i.e., requiring the reader to reach some defined acceptable high level of comprehension per unit of time with larger values deemed more efficient. This metric can be expressed as units of words-per-minute (WPM) or words-per-second (Chapman & Cook, 1923). As this measure is analogues to the conventional concepts of speed and accuracy, so too is the expectation of a performance trade-off between them under sub-optimal conditions (Dyson & Haselgrove, 2001; Wickens, Lee, Liu, & Becker 2004). As a measure of performance, reading comprehension and rate must be considered together to derive an acceptable level of confidence in any determination of reading efficiency. A composite of these measures has been used to represent a global measure of reading efficiency (Castelhano & Muter, 2001).

Although eye fixation movements, commonly referred to as saccades, typically progress across the text, regressions can occur where the eye revisits text which was previously fixated. Eye movement measurement has been utilized to detect regressions shifts within sentences which, according to Braze, Shankweiler, Ni, and Conway Palumbo (2002), are indicative of cognitive processing. Differences in saccadic behavior were used to differentiate between language syntax (form) and meaning (context). Also, it was shown that normal readers will make regressive eye movements after about every fifth word in a sentence (Rayner, 1998; Braze et al., 2002). There are also several reader eye movement behaviors that are predicable, consistent, and robust. In a comprehensive review of two decades of eye movement in reading and information processing, Rayner (1998) points out that the following phenomena are quite invariant for text reading tasks: between saccades, the eyes remain relatively still (although nystagmus is constantly present) during a fixation period of 200-300 ms. The duration of the fixation is somewhat task dependent. For silent reading, mean fixation duration is 225 ms with a mean saccade length being 2 degrees or about 8 letters at a typical font size and range (visual angle). Saccadic eye movement velocities have been recorded on the order of 500 degrees per second. As such, stimuli are not perceptible during these eye movements. A perceptual mechanism known as saccadic suppression works to keep the reader from experiencing what would otherwise be “seen” as a noisy blur during high velocity eye movements. In general, regarding readers of the English language alphabetical orthography: the size of the perceptual scan is relatively small and consistent (Rayner, 1998). The span extends from the beginning of the current fixated word but no more than 3-4 letters to the left of the fixation and to about 14-15 letter spaces to the right of fixation. Thus, the span is asymmetric and biased to the right (in the direction of typical line reading progression in Western culture). Of course, perceptual span includes stimuli sensed in the parafoveal region of the retina (5+ degrees subtended visual angle). The “word identification span” of the reading perceptual process does not exceed 7-8 letter spaces to the right of the fixation point. In summary, a single fixation between saccadic eye movements lends an identification level resolution of about 10-13 letter spaces of text.

In a series of experiments performed in an attempt to define optimal typographical factors such as text font size (Paterson & Tinker, 1929) and sentence line length (Tinker & Paterson, 1929; Paterson & Tinker, 1940), the authors demonstrated reading efficiency decrements at non-optimal dimensions. It was suggested that the saccadic like eye movement required to shift from the end of one line of text to the next was perhaps at least partially responsible for reading efficiency decrement when line length was other than optimal due either to shifts interfering with saccade rate and rhythm within short lines or next-line fixation inaccuracy shifting between long lines (Tinker & Paterson, 1929; Paterson & Tinker, 1940).

Constraining the Spatial Dimension

Reading effectiveness from a spatial optimization perspective involves the necessity of restricting saccadic eye movement to a small area or eliminating the need for saccadic eye movements altogether. A large body of research has been accomplished in this area during the investigation of a concept referred to as Rapid Serial Visual Presentation (RSVP). RSVP was originally conceived as a method of studying language processing (Forster, 1970; Potter, 1984) but the technique has been proposed as a method of providing text for readers within a small spatial area. The original theoretical basis for RSVP promoting reading efficiency is that the time required for saccadic eye movements can be eliminated by rapidly presenting a single word of a passage at a time, in sequence, at a static fixation point (Broadbent & Broadbent, 1987; Rahman & Muter, 1999). In the scientific literature, RSVP is likely the closest text presentation technique that approximates the goal of providing display of the most amount of usable text-based information in the least amount of space (display real estate). The following section provides a review of past investigations of the RSVP concept, including interaction approaches, empirical comparisons between conventional page-based reading, and competing dynamic presentations.

Of particular interest here is research where the application of RSVP was evaluated for its potential to support effective reading from small electronic display screens. It should be noted that several important questions are absent from the body of literature when considering applications and the present design problem. No work was discovered where the drawing area of the text was free to move within the boundaries of a display. This functionality may be required where the text dialog box is a label attached to a dynamic entity within large dataset visualizations (multiple interacting objects or nodes). Also absent is a record of efforts to study “drill down” text interface techniques to allow a reader to skip directly to desired bits of information known to reside within a larger set of information presented in a standardized sequence. Similarly, only a small amount of work has been done where the rapidly presented information is actually varied due to the passage of time or change-of-state. Here, there is interest in information depiction techniques that support monitoring a number of dynamic entities and affording access to associated text-based data (information). Following a review of the existing RSVP work, descriptions of other presentation and interaction concepts that may satisfy these added dimensions are introduced.

Rapid Serial Visual Presentation (RSVP) of Textual Information

Forster (1970) performed experiments to determine whether or not syntactic complexity affected the visual perception of rapidly presented word sequences. Following work which indicated that a human’s ability to organize an input sequence of rapidly presented letters was poor, a follow-on objective was to see if the same was true when the presented material had an underlying structure. In this case, the structure was provided in the form of complete sentences rather than words from a random string. It was further hypothesized that complex structure would be more difficult for observers to recall when compared to more simple structure. Although Forster’s stated objective was to determine if RSVP could be used to detect sentence complexity, the series of studies simultaneously provide a methodology for evaluating RSVP reading efficiency. For instance, three different types of presentation were constructed: simple sentences, complex sentences, and random sequencing of the words in the simple or complex conditions. In all cases, six words were presented to form a string. Words were presented to participants at a rate of 16 words per second (62.5 ms each) and each word was projected onto the retina filling approximately the same area. A film and projection technique was used to form the stimulus. In a written recall task, scored based on the correct words reported in the correct sequence, significantly fewer words were recalled when sentence structure was complex. Performance for the random word sequence was the worst, and there was no difference between the randomly presented words regardless of their origin as a simple or complex sentence. According to the author's interpretation, the syntax of the sentence structure could be detected via RSVP performance. Another interpretation is that reading processes are involved in the superior ability of the human to process text formed via an underlying structure even when presentation rate was quite high. Of course missing from this treatment is a consideration of comprehension as a performance measure, although, participants reported that they often "knew" what the sentence was about but they were unable to report actually seeing the words. This phenomenon, likely due to masking interference and/or transmission interruptions between visual storage and short-term memory (STM), is investigated in other work (Broadbent & Broadbent, 1987). Another interesting finding is that participants often reported the words from the random sequences in a way that attempted to make the words more syntactically meaningful. This research provides strong evidence that context and meaning are important information organizing factors.

Broadbent and Broadbent (1987) used a variant of the RSVP methodology to investigate a phenomenon where target word identification performance suffered from interference under conditions when the target was defined in some way that did not specify its full identity. This is analogous to a task where a user is interested in knowing that some change in the information has occurred and is further interested in being able to identify the nature of that change. In a series of experiments, participants were asked not to merely detect a target, but to also identify it. In a first experiment, the objective was to replicate a finding which indicated that participants failed to identify RSVP presented target words when the words were separated by a non-target word. Once again, this research suggests some type of masking interference likely affected target identification. Participants were made aware that there were two targets in a search list (12 5-letter words (nouns)). This was done to increase expectancy during the search and identification task. Target words were in uppercase letters vs. lowercase character distractor words. Participants were asked to write down the words after exposure to report detection and word identification. Target words appeared either together as a pair or with one, two, or three intervening distractors (80 ms exposure time per word). The overall probability of both words being correctly identified was low: $p = 0.062$. For 17 of 20 participants, the identification of the second target was worse when the first target was correctly identified. Apparently, whichever of the two targets was encoded first gained an advantage.

A follow-on experiment investigated performance when the target words were related. Again, performance for reporting both targets correctly was quite poor ($p =$

0.0625). It was also the case that reporting one target impaired the reporting of the other. This research suggests there was no effect of association.

A third experiment used longer lists of words compared to the previous study (24 5-letter nouns). In this case, non-target separation ranged from 1 to 16 words. The first target was always in the first half of the list. Targets were all presented in lowercase but designated by a hyphen on each side of the target word. At long durations of stimulus onset asynchrony (SOA), the probability of identifying both targets was almost the same as that for identification of individual targets. This research suggests that the masking was absent given this longer delay. Interference of targets lasted around 500 ms or less.

The same task was used in a fourth experiment that used different types of detection criteria without participants' knowledge of specifically which criteria they were looking for. Animal names (lowercase) were used as targets to ensure that non-targets were correctly rejected. Participants were told that each list included a pair of animal names, two uppercase letter words, a combination, or a single target of one of the two types. Again, a preceding target was disruptive to a following target. This was particularly true when an uppercase word target was followed by another uppercase word. When both target words were lowercase animal names, target reporting performance for the second word was much better.

The Broadbent and Broadbent (1987) research is of particular interest here for a number of reasons. First, compared to Forster (1970), a computer and cathode ray tube (CRT) monitor was used for stimulus generation to allow a well-controlled manipulation of SOA within the various individual experiments. Also, it is apparent that masking effects for target detection and identification tasks are somehow related to not only the temporal proximity of targets, but also to the semantic nature of the target. The fact that animal name targets (targets that were somewhat meaningful and were recognized as such based on that categorization context) seemed to interfere with each other is perhaps indicative that top-down processing mechanisms may assist transference of visual stimulus to at least working memory or STM.

It remains that this work is not directly applicable to the task of present interest because “reading” is quite different from search, detection, and identification within a list of otherwise meaningless words. That being said, this as foundational work is still important. From this research, it is understood that caution should be exercised when considering single word RSVP as a presentation technique because of the potential effect of SOA duration within sentence segment presentation. This and related studies indicate that a sort of “attentional blink” occurs that may influence readers’ ability to perceive down-stream information based on the information that precedes it during a sequential presentation (Nieuwenstein & Potter, 2006). The following discussion focuses more on the RSVP text presentation technique applied to complete reading tasks.

Castelhano and Muter (2001) described research performed to investigate the best presentation techniques for the display of text on small electronic screens. Some obvious analogies can be drawn between display of text on small a screen and the display of text within a small drawing area regardless of the overall display screen size. The 2001 study was a continuation of work performed previously by Kang and Muter (1989) and Rahman and Muter (1999). Here, RSVP was compared to several other presentation techniques such as a moving window display, “times square” (right to left horizontal scrolling), line-stepping, and sentence by sentence presentation. Because of some of the negative RSVP effects seen in previous studies, such as described by Broadbent and Broadbent (1987), the authors attempted to tune the text presentation to be more compatible with natural human reading behavior. Again, it is stated that the efficiency gain to be realized by RSVP is likely due to time savings related to the reduction of saccadic eye movements (and associated cognitive load) within and between lines of text compared to that required during reading a conventional page layout.

It was theorized that presentation rate based on word length, context significance, and sentence structure may enhance reader performance and acceptance. When the results of this research is examined from a more subjective preference perspective, while RSVP had much application potential (Muter, 1996), it was extremely disliked by readers compared to times square (Kang and Muter, 1989) and other more conventional reading methods (Castelhano & Muter, 2001). Castelhano and Muter (2001) attempted to address the RSVP acceptance difficulties through several modification methods: sentence oriented processing, addition of a completion meter, shorter presentation duration for common words, punctuation pauses, proposition integration, and interruption pauses. RSVP, in conjunction with a completion meter (which allowed readers to keep track of their position within sentences and relative to a full page layout format), common word duration, and punctuation pauses, was evaluated against normal page and sentence-by-sentence text presentations. In a first experiment, participants read passages from the comprehension portion of the Graduate Record Exam (GRE) using 12 point “Times New

Roman” font. Text was presented via an 800 x 600 pixel resolution CRT monitor. Normal page presentation was 6.5 inches across for passages ranging between 136 and 173 WPM (a mean of 153.3 words). RSVP was centered on the screen and the within sentence presentation rate of approximately 260 WPM was used (presentation rate varied slightly depending on the form of RSVP modification). A limited amount of user interaction with the text was also provided. Participants used a key to advance to the next sentence in both the sentence-by-sentence and RSVP conditions. Within the RSVP condition, another key could be used to “restart” the present sentence back to the beginning of the sentence. Previous sentences were not available. Three different conditions of RSVP were included: 1) the addition of a completion meter; 2) shortened common word duration; and 3) added punctuation pauses within the sentences. All combinations of these modifications were included in the evaluation. Performance measurements in the form of efficiency (reading rate and comprehension) as well as a reader preference scale were collected. According to the findings, normal page and sentence-by-sentence presentations were preferred over all of the other presentation techniques. It was indicated however that RSVP preference ratings increased (more acceptability) with exposure. According to performance, there were no differences among reading speed measures. Furthermore, sentence-by-sentence presentation was as well liked as the normal page format.

A second experiment looked at the use of a modified RSVP version with 460 ms punctuation pauses and interactive interruption functionality compared to a normal non-modified version of RSVP and full page presentation. In terms of preference, the modified RSVP condition was liked more than the unmodified RSVP. The full page presentation was still the most preferred reading method. There was an apparent speed accuracy trade-off related to use of the modified RSVP. Namely, the modified condition was the slowest, but accuracy measured through a comprehension test was as good as that with the normal page presentation.

Of particular interest in the studies reviewed above is the evidence they lend to the idea that appropriate modification of presentation techniques has the potential to affect both performance and user acceptance. The optimization of user acceptance is an important way to ensure that the user is not frustrated by the information seeking tasks being considered for complex data visualization applications. The existence of a speed accuracy trade-off from a performance perspective is not necessarily a condemnation of a specific presentation condition; what matters is which dimension is most desired for a specific task or application. For decision making during critical task scenarios, a design which biases responses toward accuracy over speed may be highly preferential. Another important takeaway from the literature presented here is that user interaction may play an important role with regard to the optimization of text-based information access where display space is very limited (spatially-constrained).

RSVP with Pictorial Information

Some specific work aimed at investigating spatial-temporal trade-off has looked at the rapid presentation of pictorial sequences versus textual information. Intraub (1980) performed RSVP experiments using pictures as stimuli during recognition and identification tasks. These studies found that recall after brief picture exposure was surprisingly good provided an adequate inter-stimulus interval (ISI) enabled the image to be processed by the observer. In a follow-on study, Intraub (1981) found an interaction between observer recall/identification performance and verbally cued priming. This is perhaps evidence that textual and graphical information is comparable in a way that these findings generalize to similar top-down information processes involved in text interpretation. Spence (2002) presented an information visualization technique borrowed from RSVP where the objective was to support user browsing efficiency during search tasks. Here, Spence (2002) furthered the work of De Bruijn and Spence (2000). They used the metaphor of riffling through a book to gain a general feel for its contents as an explanation of the functionality a user may gain via rapid sequential visual exposure to computer file system contents via a small screen. Similar to work presented above, it was found that superior performance was produced when observers have some interactive control of the presentation sequence and that an indicator of cycle progress is important for users' acceptance of the presentation technique. Spence (2002) defines the riffling metaphor as a means of allowing space to be traded for time with the potential to support electronic search and browsing, particularly on small displays. They point out that temporal resolution of the human visual system is limited relative to spatial resolution. This can be used to an advantage when fusion is desirable (i.e. the illusion of seamless motion during movie viewing). Some amount of fusion will occur during rapid enough presentation of imagery but, if that imagery is not properly related, the images will blend in a way that is disruptive and counterproductive. From this it can be further concluded that RSVP utility optimization is task dependent, it has significant limitations, and user control via some form of interactivity is advantageous.

Rapid Communication Display (RAP-COM)

Another area of research applicable to the present research was called Rapid Communication Display or RAP-COM by its developers (Matin & Boff, 1988; Osgood, Boff, & Donovan, 1988; Payne and Lange, 1991). Of particular interest is the hybrid nature RAP-COM represents between text-based stimulus, continuously sequenced information dimensions, and value variation within the displayed dimensions. RAP-COM frames contain chunks of text-based information similar to those that would normally be displayed on individual dials or in separate locations within an integrated display (e.g., airspeed, altitude, and heading within an aircraft display). The RAP-COM concept relies on information being related in meaning or significance. Information communication duration should be preserved via the RAP-COM concept through the reduction of the eye-movement saccade and fixation cycle requirement in a conventional display scan (225-250 ms). According to the authors, only 50 ms of that cycle is required for information acquisition. This 40-50 ms information acquisition duration is consistent with findings presented by Rayner (1998). The intent of these studies was also to use an independent measure that would be useful for comparing results across experiments with a variety of tasks. The authors defined duration threshold as exposure time given 90% accuracy task performance. In a first experiment, 0.25 degree subtended visual angle characters presented on a dark background monitor were viewed from a fixed distance by using a headrest apparatus. Using duration threshold, serial presentation was significantly shorter than simultaneous information presentation. That is when the same information was presented simultaneously but in close yet separate locations. A follow-on experiment investigated simultaneous presentation at varying display separation eccentricities (2 - 7 degrees). Duration threshold increased with separation in a linear fashion at larger separations implying that eye movement incurred a cost.

Small Screen Presentation Techniques

From all the literature presented thus far, an interesting observation can be made. In the years since the development and evaluation of the previous research, the use of small screen technology has been widely embraced as illustrated by the adoption of pocket-sized smart cellular telephones, and text presentation on these small displays depends more or less on conventional page-based layout. Although the RSVP techniques showed some usability performance potential, there is little evidence that they found any practical application beyond some value as a research tool. Instead, higher pixel density displays viewed from shorter distances facilitated the presentation of text which required multiple fixations per line of text.

Of course some dynamic text presentation techniques are quite common. Öquist (2006) performed a series of experiments with the aim of finding the best way to present text-based information on screen sizes appropriate for mobile devices such as personal digital assistants (PDAs) and mobile telephones (presently, a combination of these technologies exists as the “smart phone”). Scrolling, leading, paging techniques, and RSVP were investigated through the use of eye tracking measurement and task loading index ratings.

Scrolling presents text in the traditional page format within a display where the page is larger than the screen drawing area (only part of the page text area can be displayed at a time). Interactive scroll bars are used to indicate how much of the text is being displayed on the screen as well as the current horizontal and vertical position within the larger text area. For small screen presentation, the width of the text strings is often narrowed to fit within the horizontal allowance of the screen. In this way, the user need only use a scroll bar in the vertical dimension to gain access to the full length of the passage. The user is afforded navigation through the text by interacting with vertical movement of text. For the Öquist (2006) evaluations, this interaction was accomplished through up and down button control for single line advancement or reversion. Present technology more often utilizes touch screen dragging with physical law-based dynamic momentum characteristics for scrolling-type manipulation. For large screen presentation of complex data set visualizations, it is likely that a mouse or trackball pointing device would be the interaction control technology of choice. Of course, depending on the task, touch screen and/or gesture recognition approaches should be considered.

Paging presents all the text that can be fit within the vertical and horizontal dimensions of the display as a discrete partial “block” or “chunk” of the overall passage. Via some input control device, the user may navigate forward or backward through the passage a “page” at a time. In support of this approach on small screens, a location meter of some sort is typically needed to orient the reader within the passage. The least number of interactions required for a user to see all of the text is equal to the number of pages required to display all of the text chunks. Compared to scrolling, the number of interactions required of the user would be expected to be much less.

Leading (which is referred to from this point forward as times square (TS) presentation) is a technique where one line of text dynamically scrolls smoothly across the screen from right to left. Recently this has been referred to as the “ticker” or the “crawl” of text information across the screen during a news program on television. Roth (2008) investigated the potential to scroll diagonally across a screen to maximize string line length and font size of the presented text. Regardless, it seems some form of completion meter was needed to help a reader to know how much of the passage remains to be fed across the screen. Interaction options include control of the scroll rate (typically in WPM units), starting, stopping (pause), as well as forward and backward movement functionalities. TS is a candidate for complex dataset and multiple entity text labeling where the drawing area is severely limited both vertically and horizontally. Öquist did not consider the vertical corollary to the leading technique. This approach has the text dynamically scroll from the bottom to the top of the display. This technique is referred to here as cinema credits (CC) due to its similarity to the way credits scroll from bottom to top of the screen at the end of a feature film. Of course a major consideration for the CC approach is the available horizontal and vertical display drawing area.

Readability evaluations of these various techniques by Öquist (2006) found that, similar to previous studies, RSVP was not as well-liked by users as other presentation formats. This was true even when reading efficiency was the same or better than other presentation techniques. In a series of studies where eye movements were measured, the theory that RSVP gains in efficiency (increased reading speed and reduced cognitive load) by reducing or eliminating eye movements was not supported. In the studies performed by Öquist (2006), RSVP was far less efficient compared to the other formats. Öquist eloquently states: "A new text presentation format does not really let you do new things; it lets you read things in a new way. For a new text presentation format to rival existing formats it must probably offer immediate gratification, either in terms of increased readability or something else." Going forward, this "something else" Öquist refers to may be a functionality for a user which affords fairly easy access to text-based information in a way that does not include an associated trade-off of potentially disruptive occlusion related to the presentation of multiple lines of text within a conventional format width such as one that would approximate the horizontal display area required to present most of a complete sentence.

Summary

The objective of this chapter was to review the relevant literature related to the effective display of text-based information within small drawing areas. A body of work was found that was helpful toward answering some of the basic perceptual issues related to text readability and the display of information via electronic and small screens. Interestingly, very few of the techniques discussed in the past literature have been adopted into mainstream use even in the wake of the proliferation of small screen devices. Nonetheless, the findings of these efforts provided generalizable measures and methodologies to support the direction of the research presented here.

No literature was found that directly applied to the presentation of relatively large amounts of text in a small area to support complex dataset visualizations. More specifically, there was a lack of information about designs intended to support C^3I-type tasks and their associated demands for decision makers. For this application, the display surface itself may be quite large, but text is co-located with some number of dynamic and interacting entities. To address this, new and different presentation techniques and user interaction mechanizations were developed and evaluated. As is seen in most of the previous literature, an undertaking like this was expected to involve performance trade-offs in multiple dimensions. For that reason, well-designed studies and an iterative design process needed to be pursued to conclude with any confidence (and generalizability) that a performance enhancing solution may have been (or can be) achieved. At the very least, it is important that an appropriate use case scenario for an applied evaluation of candidate presentation and interaction concepts be developed and utilized. Similarly, the development of a valid and reliable evaluation methodology was critical. The evaluation methodology took into consideration whether the scenario has a logical primary/secondary task prioritization. The evaluation was a scenario designed so that it is based on elements common to the real-world applications of interest. Done effectively, keeping those elements representative of the real-world was intended help ensure that the evaluation scenario and associated performance tasks were internally and externally valid. Critical to the value of the methodology was the identification and implementation of potentially sensitive performance measures to form the basis of a practical performance comparison of the human-system interface variations of interest. Here too, looking back to the body of previously accomplished work has helped to inform how the present research was conducted.

The following are contributions uncovered by the previous research which were built upon for the research presented here:

1) A single fixation between saccadic eye movements lends an identification level resolution of about 10-13 letter spaces of text for a typical font size and viewing distance (Rayner, 1998). This seemed to define an upper boundary for how large a single fixation display constraint should be.
2) Research showing the potential for RSVP was consistent when results originally collected by film and projection were replicated using electronic displays (Broadbent & Broadbent, 1987; Forster, 1970).
3) Studies performed to "tune" text presentation for human reading behavior, such as saccadic eye movement, showed performance benefits (Castelhano & Muter, 2001; Kang & Muter, 1989; Rahman & Muter 1999). This indicated that a similar approach should be used when text is in full sentence form.
4) The addition of progress monitoring and direct user control of spatially-constrained text presentation showed promise toward both enhancing performance and user acceptance (Castelhano & Muter, 2001; Spence, 2002; Öquist, 2006). This motivated the use of similar features for the research presented here.
5) Rapid presentation of information when changing values for system parameters where presented versus natural language text showed performance benefits (Matin & Boff, 1988; Osgood, Boff, & Donovan, 1988; Payne & Lange,

1991). This category of information is most similar to that employed for this research versus the context contained within the presentation of complete sentences.

III. Experiment 1

Secondary Task Difficulty Determination Study

Chapter Overview

This research uses a secondary task paradigm as introduced in Chapter I. A central concept of the secondary task paradigm is that performance measures of the secondary task be sensitive enough to detect differences among the independent variable manipulation of the primary task. As such, it is critical that the “difficulty” of performing the secondary be properly developed so that it is challenging to perform yet near perfect performance is achievable given the exertion of reasonable effort. This chapter presents a study that was conducted to determine the appropriate level of difficulty for the secondary task to be used during the overall experiment (Experiment 2) toward the objective of evaluating the text access interface variations. Experiment 1 was conducted and prepared as a standalone manuscript for publication submission to the journal *Theoretical Issues in Ergonomics Science*. At the time of this writing, the submission has been accepted for publication. The submission manuscript is presented in its entirety below. By publisher request, the manuscript was prepared using United Kingdom English spelling. The figure and table caption numbers have been modified to be consistent with this document.

Theoretical Issues in Ergonomics Science Paper

Development and validation of a secondary task environment for assessing visual-psychomotor tasks

The dual-task paradigm methodology is a widely accepted approach to facilitate relative human performance measurement in a variety of tasks. The present paper describes the development and experimental validation of a visual-psychomotor secondary task. This task is proposed as a standardised secondary task set to facilitate human performance measurement during the objective evaluation of alternate primary tasks. The development of the secondary task is aligned with attributes suggested within the existing literature. The methodology is offered as a systematically derived secondary task with a tuned difficulty level which is intended to avoid floor or ceiling effects. Also, the data presented here afford the reader the ability to manipulate difficulty with known effect, if desired. Future plans include use of the secondary task set to facilitate a comparison of primary task independent variables. This activity will act to further exercise the potential utility of this 'standardised' secondary task and its associated mechanisation.

Keywords: Human performance; dual-task paradigm; secondary task; visual information processing; standardised task.

1. Introduction

The dual-task paradigm is a well-established experimental methodology for human performance evaluation. The method typically involves the manipulation of one or more independent variables within a primary task while a decrement in human performance is measured in a common secondary task to provide an objective, unbiased comparison. In this paradigm, it is not necessary to observe differences in primary task performance to make quantitative inferences regarding the cognitive demands that manipulation of the independent variables place on the human user.

If one assumes that the human can be modelled as having a single, limited pool of cognitive resources that must be shared by all tasks, as assumed by Kahneman's Capacity Model (Kahneman, 1973), one can provide a very simple and logical explanation of the dual task paradigm for a capacity limited task (Kantowitz & Sorkin, 1983). In this explanation, it is assumed that humans have a limited pool of cognitive resources that constrain their capacity to process information and perform tasks. It is assumed that these resources are adequate to permit the user to perform the primary tasks. It is also assumed that when there is a requirement that both the primary and secondary tasks be performed concurrently, humans can perform the tasks without decrement provided the 'pool' of resources available for the effort is not exceeded by the difficulty of the effort. However, if the effort to perform the secondary task simultaneously with the primary task exceeds the available resources, some performance decrement across the tasks is expected. If one of the tasks is prioritised by the human as primary, it follows that any performance decrement present will be isolated to the secondary task. As a result, any significant decrements in performance of a secondary task while performing one primary task as opposed to another can be interpreted to occur as a result of different levels of cognitive effort or 'mental workload' due to the primary task manipulation. If we equate this change in level of cognitive effort to changes in 'usability' existing among the levels of the independent variables, then we can measure the changes in the usability of the system employed to facilitate the primary task by measuring changes in secondary task performance as a function of the independent variables.

Unfortunately, application of this method becomes more complicated as we apply more modern models of cognitive resources, such as Wickens's Multiple Resource Theory (Wickens, 2008). The implication of this model is that tasks that are more successfully time-shared will compete less for common resources than tasks that cannot be performed concurrently with ease (Kramer, Wickens, Donchin, & Fregly, 1985). Therefore, these tasks will make less effective primary-secondary task parings. This implies that an effective application of the dual-task paradigm should be limited to the same resources for both the primary and secondary tasks.

Efforts have been made to define the properties that an appropriate secondary task should include for successful use. For example, Fisk and colleagues outline three rules that a secondary task should satisfy to be useful: (1) the secondary task must remain resource sensitive throughout the experiment; (2) there should be equivalence of single and dual primary task performance; and (3) there should be a resource trade-off with the secondary task which is sensitive to the resource demands of the primary task (Fisk, Schneider, & Derrick, 1979).

Although the dual task paradigm is well recognised, a standardised set of secondary tasks has not been defined to date. In the current research, we seek to develop and validate a secondary task that can be applied in conjunction with primary tasks that involve the assessment of visual information paired with psychomotor tasks. Methods such as the well-known Multi-Attribute Task Battery (MAT-B) and its variants (Comstock, Jr. & Arnegard, 1992) represent examples of standardised benchmark operator performance evaluation tools. This paper is organised to describe a proposed secondary task and associated performance measures, as well as a method to validate the proposed task.

2 Materials and Methods

2.1 Defining the Secondary Task

The overall goal of the present research is to define a secondary task which is sensitive to changes in a primary task which requires the visual and psychomotor channels.

2.1.1 Secondary Task Requirements

Based on the literature and the need to develop a relatively universal task, the task shall fulfil the following criteria:

1) require constant visual attention to be successfully completed;
2) provide measureable results that are sensitive to lapses in visual attention;
3) permit task difficulty to be adjusted over a large range;
4) provide multiple conditions with equivalent task load;
5) be easy to learn, providing no more than 20 minutes of training;
6) not require special skills or abilities beyond basic visual performance and the ability to perform psychomotor tasks common in interaction with a desktop computer; and
7) easy to replicate without special hardware or software.

A secondary task is designed to address these criteria.

2.1.2 Secondary Task Description

The task requires a participant to monitor the activity of several entities on a visual display and report observed entity status changes as they occur. Figure 4 depicts the environment containing three entities. During a trial, the entities move across the screen at a constant rate in one of four possible directions relative to screen coordinates: up (+Y), down (-Y), right (+X), or left (-X). Entity identification is differentiated through the use of differently shaped and colour coded symbols: green circles, red triangles, and yellow squares. Each will include its associated entity member number within the centre area of the entity shape. Attached to each symbol is a data tag intended to provide specific identification labelling for the associated entity.

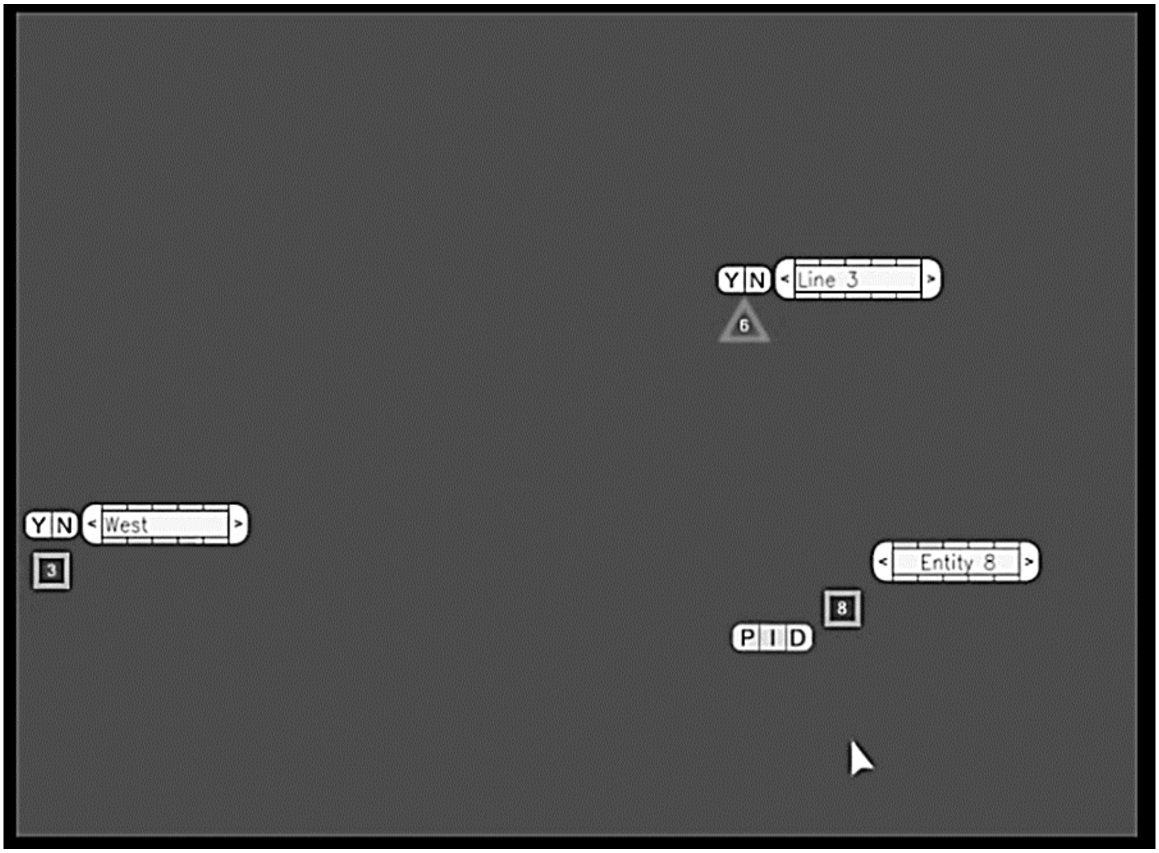


Figure 4. Multiple entity identification symbols, label tags, cursor, and response box.

Participants are instructed to monitor the multiple entities for three specific events that are 'reportable' within the scenario. Table 1 shows the three different reportable entity events. The following paragraphs will describe how the secondary task events are scheduled and how the participants interact with the secondary task.

Table 1. Reportable entity events for the secondary task.

Entity Change	Description
Presence (P)	entrance of entities into the scenario from 'off screen' as they become visible
Identification (I)	the current entity changes shape and colour to become one of the other two types of entities
Direction (D)	the motion of an entity changes from its current trajectory by a +/-90 degree or -180 degree change from its current course

2.1.3 Task Event Control

According to a controlled schedule, entities enter and exit the scenario. No change events occur during the first 15 s of runtime to allow the scenario to be fully populated before interaction beyond reporting presence is required of the participant. After that point in time, the scheduling logic loops through all of the active entities that are not scheduled to exit the scenario within 10 s given their current trajectory, or have had the same event occur within the preceding 10 s. If no entities meet the event criteria, a ‘blank’ event is recorded. For an eligible entity, an event is assigned according to the following probabilities: direction change (p = .40), identification change (p = .40), and no event assignment (p = .20). The frequency of events is controlled by a min/max event timer and is thus controllable with some precision. Similarly, the scenario is designed to maintain a fairly constant mean number of active entities at any given time once the general population is established (at about the initial 15 s point). A logical schedule is employed here as well. If the scenario is underpopulated by one, a new entity will be added according to a uniform random distribution between 10 and 20 s. If the scenario is underpopulated by two or more and the next entity is not scheduled to enter for more than 10 s, then an entity is scheduled to enter at a random time between 5 and 10 s. If the scenario logic detects that two entities will exit the scenario within the next 10 s, and no new entities are scheduled to enter within the next 10 s, a new entity is scheduled to enter the scenario at a random time between 2 and 4 s.

2.1.4 Secondary Task SysML Activity Diagram

Figure 5 depicts the activity elements participants will encounter while performing the secondary task. The initialisation point in the diagram represents a point that a 600 s trial is underway. The total session length can be any duration that is appropriate for the desired primary task (although sessions less than 5 minutes in duration are not recommended). In general, the secondary task requires the participant to monitor the entities and indicate any reportable events. Participants select entities and report change events by hovering the mouse cursor over the desired entity and actuating right or left mouse clicks for response input. The diagram indicates that the secondary task activity can be interrupted by a primary task requirement. Subjects should be encouraged by direction and training to prioritise performance of the primary task when the interruptions occur.

The session is terminated when there are no longer any entity events to be reported, there is no active query activity, and the endpoint has been reached (for this example 600 s).

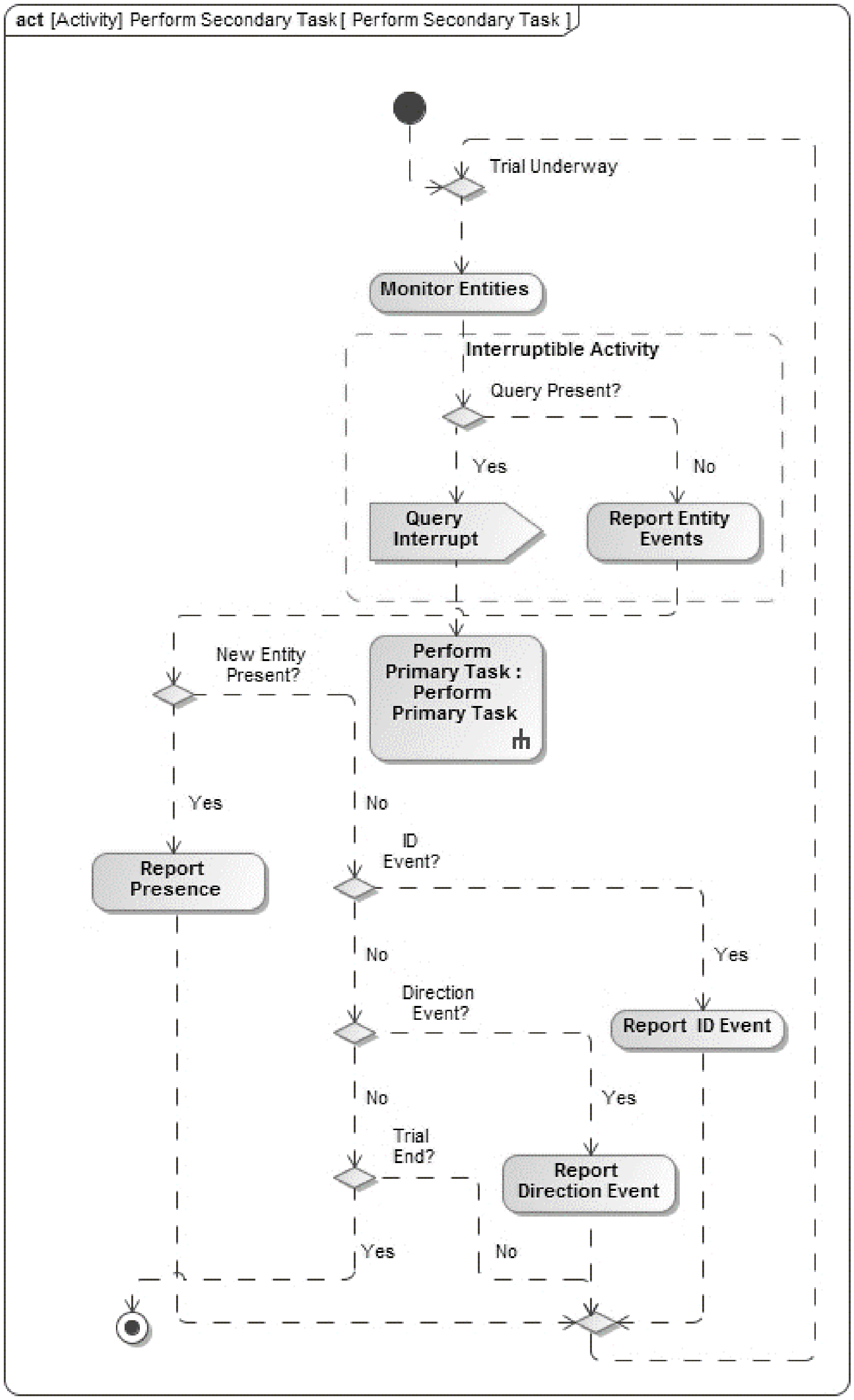


Figure 5. Secondary Task SysML Activity Diagram.

2.1.5 Secondary Task Difficulty

The intent of the event schedule logic and its associated monitoring functions described above is to generate the appearance of event randomness for the subjects. In fact, the events are highly scheduled and controlled. The desire to control event scheduling is two-fold. First, the basis for a constant set of scenario events can be presented to subjects in a way that they are not able to predict a learned sequence of events. This is accomplished by making event initialisation dependent on the interaction of the participant and the ability to initialise a common scenario with different initial conditions (e.g., first entry direction). Second, difficulty of the secondary task can be manipulated directly by controlling the average number of entities in the scenario and the rate of change assigned to the event schedule probabilities within the bounds of the control criteria described above. Of course difficulty can be manipulated by speeding up or slowing down the entity motion rate as well. It is highly desirable to tune the difficulty of the secondary task so that, when performed in isolation, it is engaging with reasonable effort but near 100% accuracy performance is achievable (Fisk et al., 1979). This forms the basis of the secondary task measure that is intended to be sensitive to the lack of spare capacity (Kramer et al., 1985) afforded to subjects given the simultaneous performance of the primary task and the manipulation of the independent variable levels within a primary task.

2.1.6 New Entity Initial Condition

When a new entity is added, the simulation logic assigns a random direction. However, the direction is constrained to enter from a different side from the previous entity, spreading the entities across the display. The entity is then assigned a random position along the border from which to enter. Entities are permitted to overlap. However, a layering scheme is used so the new entry is ‘on top’ of the otherwise interfering entity. Similarly, any selected entity is moved to the top layer of the scenario. Due to the speed of the entities and their size, an entity entering from the left or right will not be completely visible before a 10 s event response time limit expires, but presence reporting can be accomplished by selecting any visible part of an entity to report presence.

2.1.7 Subject and Task Interaction

All direct interaction between the participant and the scenario is accomplished with a standard ‘mouse’ input device. The mouse is used to move a cursor to any location within the bounds of the scenario area (defined by the outside border depicted in Figure 4). Proximity of the cursor to an entity automatically associates that entity to the cursor for purposes of selection and further interaction. A ‘right’ click of the mouse device button is used to select the entity. Once an entity is selected, an input box is presented at the tip of the cursor arrow permitting the entity change to be logged. Three buttons within the box are labelled with a ‘P’ for reporting presence of a new entity, an ‘I’ to report an identification change, and a ‘D’ to report a direction change. Once an entity is selected and the reporting buttons are present, the box remains stationary relative to screen coordinates (fixed in x, y screen coordinates position) until a right or left click is input with the cursor hovered over the P, I, or D button within the ‘PID box’. When the cursor is hovered over one of the PID buttons, the button nearest the cursor tip is highlighted. Whatever button is highlighted when a mouse button is clicked is recorded and logged as an associated input. Figure 4 depicts a static PID box that appeared when ‘Entity 8’ was selected via right click. Entity 8 has continued to move away from the location where it was selected. The location of the entity selection is approximately where the PID box is drawn. The PID box will remain visible until a P, I, or D button is selected via a mouse click (PID choice condition), a right or left click is made near another entity (re-selection condition), a right or left click is made away from any entity or PID box (cancelation condition), or 10 seconds have elapsed without any click made by the participant (timeout condition). No more than one PID box is present at any time.

2.1.8 Secondary Task-Dependent Measures

The secondary task-dependent measures are largely focused on accuracy and response time. An accurate input is defined by the selection of the proper entity and PID box button for the preceding change event. Errors are defined several ways: (1) failure to select an entity within the maximum allowable time of a PID event; (2) failure to select the correct PID event change button within the maximum allowable time for the correct entity (PID button selection must be made within the maximum allowable time of entity selection or it will be reset); (3) selection of the proper entity but selection of the incorrect PID event button; or (4) selection of an entity not associated with a preceding change event.

The following secondary task-dependent measures are recorded for purposes of findings analysis and findings interpretation:

(1) Accuracy measures: correct entity selection, incorrect entity selection, correct PID selection, incorrect PID selection, and change event misses.

(2) Response time measures: total response time per change event, elapsed time from event change to entity selection, elapsed time from entity selection to PID box appearance, and elapsed time from PID box appearance to its removal (i.e., event selection or cancellation).

(3) Timeout occurrences.

Other response measures may be of interest to ensure a thorough, interpretable, and informative analysis. For instance, cursor activity is an objective measure that is likely to co-vary with the other measures already discussed, and it may add to the ability to make inferences regarding the usefulness of the different primary task display formats. Cursor activity as a continuous measure may inform how much work the participant had to perform to achieve a level of accuracy within some response duration. This measure lends itself to investigation in terms of movement area, relative location (to other objects), and rate of input. At the end of an active cursor input associated with a change or query event (indicated by its motion and subsequent halt), the location of the cursor on the display screen may be a good approximation of instantaneous participant eye fixation. This may also be informative in terms of participant task engagement and overall effort.

In general, all events are recorded with enough frame rate resolution to allow the data collection sessions to be replayed (at system clock update rate) along with all system input actions. This will enable the data reduction to be revisited in case new and unanticipated measures are desired for analysis.

2.1.9 Task Training and Feedback

After completing a written and verbal description from the experimental tasks, participants should be trained on the primary, secondary, and combined tasks until the effects of learning on performance become asymptotic. Participants should be trained first on the secondary task procedures until performance is nearly error free and workload is acceptable. Participants should be instructed to maintain a continuous visual scan pattern so that change events among the entities are not missed. Next, participants will be familiarised with the primary task by performing it in isolation of the secondary task. Once participants are comfortable with the primary task and performance has stabilised, the primary and secondary tasks should be combined to train simultaneous performance of both the primary and secondary tasks. It should be reiterated that the participants prioritise performance on the primary task and, within both tasks, prioritise accuracy performance over response speed. Participants should be reminded to perform the secondary task to the extent possible given any spare capacity to do so.

During the training sessions, participants can be provided feedback to help inform them of their accuracy on the primary, secondary, and combined tasks. Training performance and feedback can be monitored by the experimenter so that errors can be pointed out and tips for optimal performance of the tasks can be communicated and standardised across participants. The objective is for all participants to perform every task as similarly as is practical so that their performance is representative of trained operators. This process can be monitored to ensure that all participants have the motor dexterity and skill to perform the experimental tasks.

2.1.10 Subjective Feedback Data Collection

Additional measures of interest include subjective ratings. It is important to collect subjective workload measures at appropriate points within the experiment. This measure provides a global measure of participants' perceived level of effort (based on spare capacity) which simultaneously includes primary and secondary task activity. This measure can be correlated with the objective measures. Subjective preference data should also be collected to analyse consistency or lack of consistency with the other measures. This is useful for interface comparisons where objective measures are inconclusive or inconsistent with otherwise strong participant indications of preference (Castelhano & Muter, 2001).

Although any number of workload scales might be used, during verification, participants completed a Bedford Workload Scale decision tree procedure (Roscoe & Ellis, 1990). In this procedure, participants report subjective workload, applying a rating between 1 (workload insignificant) and 10 (tasks were abandoned because the participant was unable to apply sufficient effort) based on consideration of spare capacity to perform tasks in addition to the primary and secondary tasks. Use of the Bedford Scale is described during participant introduction and instructions. A practice data collection (task training) session should include demonstrating use of the scale.

2.1.11 Experimental Control (Minimising Individual Differences)

A tendency to trade-off speed for accuracy or vice versa, is one of the individual characteristics that could vary between participants and add unwanted systematic differences to the response variables within some studies (CIE, 2002). Variability in this trade-off detracts from the ability of the experimental procedure to isolate experimental comparisons. Additionally, there are potential environmental influences on task performance which should be minimised to the extent possible to reduce systematic error.

Additional individual characteristics can occur due to differences among humans within any population. Experimental control can take the form of participation selection criteria based on appropriate combinations of physical capability (visual acuity, handedness, etc.); skillset (past experience); training; or demonstrated performance (motor dexterity and skill). When these variables are not controlled directly or where unwanted systematic variably is suspected but is unavoidable, data regarding these differences should be recorded for later analysis. Potential influential conditions that are less controllable exist as well. Examples are participant motivation, pre-existing fatigue, different task completion strategies, and differing priorities. These influences can be problematic as they can vary within individuals during the study.

In general, acceptable motivation can only be assumed in the absence of competitive performance-based compensation. For volunteer or low-value compensation participation, little control of motivation can be expected. Strategy and performance prioritisation variability (response accuracy versus response speed) is mitigated through instruction and training. Participants should be instructed to prioritise the accuracy of their responses over the speed of their responses and this prioritisation should be reinforced during training. Also, in an attempt to mitigate individual skill differences among participants, as well as ensure that they are adequately familiar with the tasks, training to a baseline level of performance should occur prior to formal data collection

(Chase, Irwin-Chase, & Sonico, 2004). Time-to-train can be recorded for analysis. The utilisation of a within-subjects experimental design is another individual difference mitigation technique. Counterbalancing should be used to avoid order and learning effects for within-subject designs. Fatigue and its influence on attention, motivation, and etc. can be mitigated by including proper rest and recovery periods between data collection sessions. A session duration that does not require unacceptable sustained effort should be utilised.

Differences within the data collection environment can be controlled where it is anticipated that unwanted systematic effects may result. Data collection should occur in the same physical location using the same lighting conditions and workstation (display and input devices). Participant posture (seated positon at a workstation display with a keyboard and mouse) may be allowed to vary to suit each individual's comfort preference. The workstation display (monitor) location should be fixed across sessions.

2.2 Experimental Validation

To validate the secondary task, a study was conducted to determine the appropriate secondary task difficulty. It is desirable that the secondary task be engaging yet, when performed in isolation, afford near errorless performance. A Bedford Workload subjective workload questionnaire was administered to ensure that self-reported workload indicated enough capacity to allow performance of the primary task while remaining challenging.

Task difficulty for this study was defined via the following formula: difficulty (d) = number of active entities (n) x event rate (er) or $d = n*er$. First-hand experience with the task as well as existing literature (Oksama & Hyona, 2004; Pylyshyn & Storm, 1988) was utilised to select the number of entity levels to be investigated for the study as an independent variable manipulation (n = 3, n = 4, and n = 5). Within each entity number level, *er* was manipulated so that the rate increased according to the scheduling logic. Participant performance was recorded. Rate of motion is another available difficulty variable, but entity velocity was held constant at 3.5 deg/s throughout the study. The objective of the study was to empirically determine the values of n and *er* which provide the desired level of d for a properly sensitive secondary task.

2.2.1 Participants

A total of 10 participants volunteered for and completed the study. Average age of the participants was 37.9 years and included one female and nine males. The mean reported video game use was 8.4 hours per week among the participants who reported video game use. The range of video game use was from 2 to 20 hours per week with a median of 8 hours use per week. Five of the participants reported not using video games. On average, computer use was reported as 39.9 hours per week while television viewing was reported as 10.8 hours per week. Video game use was not considered as group discriminator for analysis due to the highly confounding nature of video game use and participant age (as the median age for game users was 28 while the median age for non-users was 48).

Prior to participation, participants received an email containing a detailed description of the study, an outline of the minimum vision performance requirements, and the expected duration required for participation (a 30 minute training session followed by nine 10 minute data collection sessions). Participants completed the data collection sessions across multiple days with the stipulation that they complete any 10 minute session once it was initiated. Participants reported normal or corrected-to-normal visual acuity and normal colour vision to participate in the study. Only two of the 10 participants reported corrected visual acuity and in both cases, vison correction was by prescription spectacles. Two participants reported left handedness while the rest indicated right handedness.

2.2.2 Apparatus

A 43 cm LCD monitor (2 megapixel resolution, 4:3 aspect ratio) was utilised to present the visual stimulus. Participants were seated at a viewing distance of approximately 60 cm. A conventional computer processing unit equipped with a standard 101 key keyboard and mouse was used to generate the stimulus display and collect participant response input, as well as workload ratings. Demographic questionnaire responses were collected using pen and paper.

Software development included generation of all the visual elements, task mechanisation, performance data collection routines, raw data recording, and initial data reduction. C++ and wrapper code for the secondary task set as presented here is available by contacting the second author directly.

2.2.3 Data Collection Procedure and Sequence

The experiment consisted of the following sequence:

(1) *Introduction, safety briefing, and consent*. Participants read a short description of the study rationale and sequence of events. Participants read an informed consent form and signed it to confirm their agreement to participate in the study. Participants were

informed that their participation in the study is completely voluntary and that they could withdraw their participation for any reason.

(2) *Demographic questionnaire.* Participants completed a questionnaire to report past experience with video displays and tasks which are similar to those used in the study.

(3) *Vision screening.* Participants were asked to report normal or corrected-to-normal visual acuity and normal colour vision. Corrective lens were used for all phases of participation in the study if worn normally for similar computer tasks.

(4) *Task training*. Participants were given detailed instructions regarding the task operation, objectives, and desired performance. Participants were also given practice during a fully-dynamic task session. The practice session included specific instruction for, and hands-on experience with all interactions they performed during the actual experimental events. The entire range of input and output variations were covered during this training period. In most cases, this training lasted less than 30 minutes. During the training session, participants were able to pause and ask questions concerning the tasks, although participants rarely exercised this ability. During this time, participants practiced completing the Bedford Workload Scale. Once the experimenter and participant were satisfied that the tasks were fully understood and adequate performance was demonstrated, the data collection session was initiated.

(5) *Data collection*. Each participant completed nine data collection sessions, each lasting approximately 10 minutes. Within each session, the rate at which PID events occurred was increased according to the schedule logic described earlier. Also, across the sessions, the average number of entities in the scenario was three, four, or five.

(6) *Workload assessment*. Workload reporting was performed by participants completing a Bedford Workload Scale at the completion of each of the 10 minute sessions. Participants were instructed to consider the workload rating for the previously completed 10 minute session in total versus answering the spare capacity decision tree for just the end of the session. The intent was to capture the extent to which participants reported that the task difficulty manipulation ‘felt’ different at the various levels under which they performed the task.

2.2.4 Task Difficulty Manipulation

Secondary task difficulty was manipulated by using a duration-based ramp-up technique where again, difficulty is objectively defined as: difficulty (d) = number of active entities (n) x event rate (er) or $d = n*er$. Difficulty was incrementally increased and data were collected to detect when collective participant performance was negatively impacted. During each session, task difficulty was manipulated by increasing the entity identification change frequency and increasing the frequency of entity direction changes to effect changes in event rate. Since each of these events require input interaction by the participant, it follows that the task became increasingly busy, errorless performance became increasingly unlikely, and the effort required to attain errorless performance increased. Difficulty was further manipulated by altering the average number of entities within the session that the participant was required to scan and monitor for change events ($n = 3$, $n = 4$, and $n = 5$). Each participant performed the task first with three entities, followed by four, and finally five. This sequence was designed to give participants the benefit of learning while the task became more challenging due to the addition of entities.

Event frequency was scheduled via a min/max timer. The initial value for event occurrence was 6 s (min) and 12 s (max). Every two minutes the min/max event time was decreased by 30%. This had the effect of 'ramping up' the change frequency among the entities thus increasing the *event rate* within the difficulty formula. Each of these levels are labelled as separate 'Event Rate' intervals within Table 2. When appropriate, Event Rate Interval 1 was not entered into analyses because this was the period of time (2 min) over which entities were entered into the scenario to reach the desired mean number for the appropriate trial type ($n = 3$, $n = 4$, or $n = 5$). The addition of entities during Interval 1 period can be seen in Figure 6.

Table 2. Event min/max frequency increases over time during a 600 s session.

Event Rate Interval	Time (s)	Minimum (s)	Maximum (s)
1	0-120	6.00	12.00
2	121-240	4.20	8.40
3	241-360	2.94	5.88
4	361-480	2.06	4.12
5	481-600	1.44	2.88

The scenario allowed the selection of randomisation seeds so the same sequence of occurrences resulted for any particular seed. This was desirable so that the eventual scenario events could be held constant across conditions for the secondary task (thus reducing unwanted variation while remaining unpredictable to participants). The scenario also allowed the selection to invert the x and y dimensions so that the same sequence of events could be presented during multiple trials but continue to appear random and unpredictable as transposition of these dimensions provided variation in the spatial pattern of events.

Seed selection was determined by running the scenario with three different random seeds for a 10 minute period in the absence of any input interaction during a five member active entity scenario. Figure 6 shows the result of entity presence for the three runs.

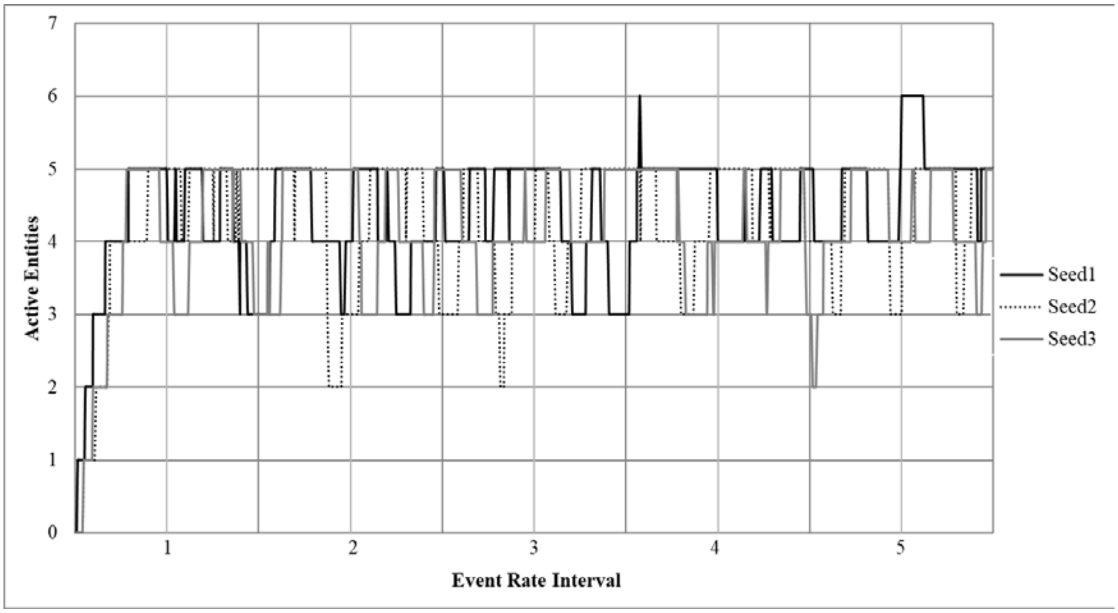


Figure 6. Seeded entity presence results via all scenario constraints.

With all constraints in place to keep the average number of active entities consistent, the result was periodic intervals where the number of active entities was reduced by one, two, or occasionally three below the desired number of entities. The random seed procedure afforded a process by which a scenario selection could be made to minimise this occurrence, but the random seed also affected entity-level change events within the scenario.

Figure 7 is a plot of the events per 30 s second window using the same scheduling logic as portrayed in the Figure 6 seeded scenarios. The plot shows the number of events per entity increasing as expected. Given the random seed procedure, it is apparent that there could be considerable differences in the number of events between different schedules. For this reason, a seed selection process was utilised to generate scenarios that best represented the desire to maintain a fixed mean number of entities in the scenario as well as a smoothly increasing *event rate* across the session period.

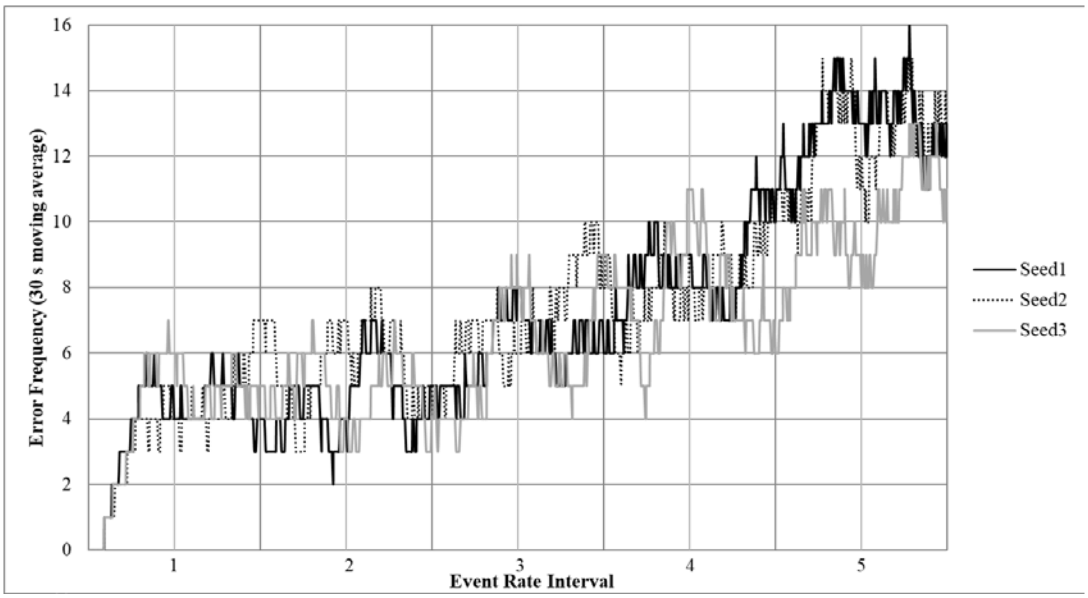


Figure 7. Seeded entity presence results via all scenario constraints.

Nine different random seeds were created for each of the three entity number settings (n = 3, n = 4, & n = 5). The sessions were run in the absence of interaction input to observe the effect of the random seeds on change event occurrences as a result of the scheduling logic described above. The resulting 30 s moving average of change events were plotted so they could be inspected for acceptable patterns of consistent *event rate* ramp-up within the scheduled two minute blocks. Figure 8 is an example plot that contains an event *rate* reversal within the 361-480 s time interval. This reversal is not desirable, and thus the random seed was rejected. Figure 9 is an example of an acceptable random seed based on the consistent stair-step characteristic of the resultant *event rate* pattern. Figure 9 is an example of a random seed that was selected for inclusion in the study.

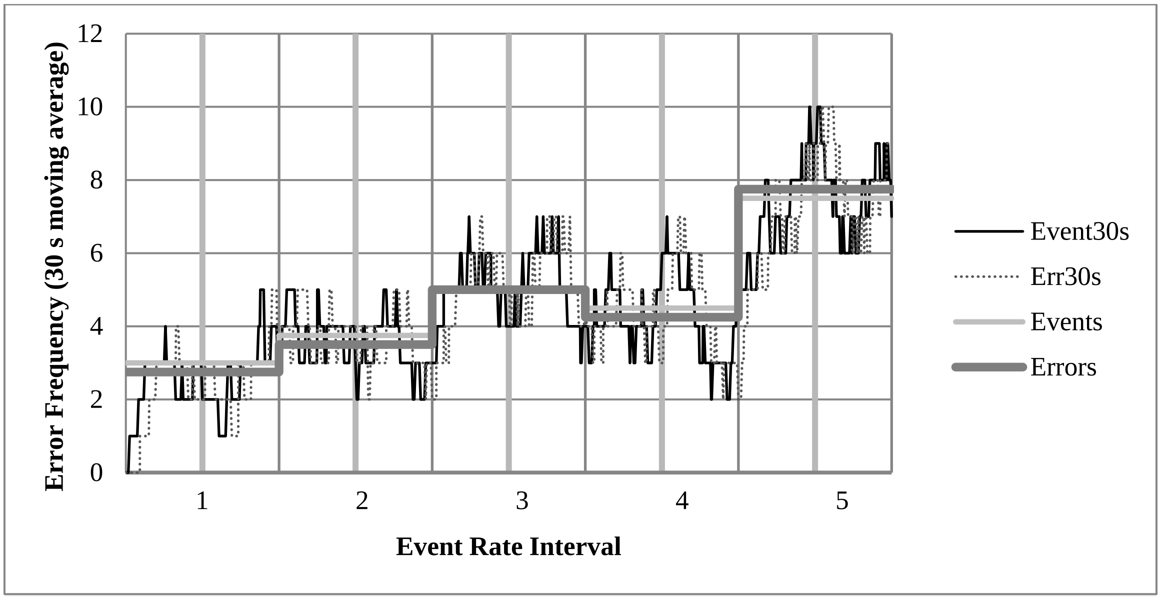


Figure 8. Change event random seed with reversal.

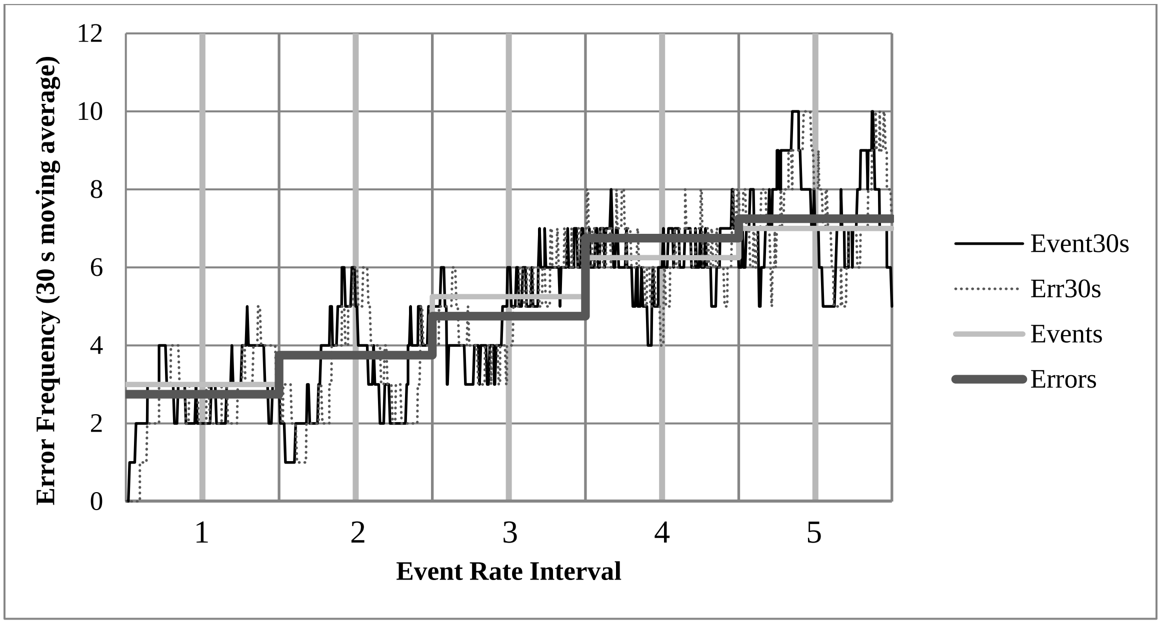


Figure 9. Change event with acceptable random seed result.

An Excel workbook was created to launch sessions and track participants' progress. Three replications of each condition were completed by each participant. Each replication differed in that one was an original *event rate* schedule and the remaining two were the result of inverting the x and y coordinate axes to produce less predictable change event occurrences. In total, this resulted in nine data collection sessions for each of the ten experimental participants.

2.2.5 Data Analysis

The data collection routine included reduction of error occurrences into 30 s moving means output, as well as calculation of errors and response times for each 2 minute block of constant event rate. Three-way, repeated-measures analyses of variance (ANOVAs) were conducted including independent variables of number of entities (i.e., 3, 4, or 5), event rate (see Table 2), and direction (i.e., (x, y), (-x, y), or (x, -y)) for the dependent measures of error and response time. A two-way, repeated-measures ANOVA was conducted to evaluate the effect of number of entities and direction on perceived workload.

3 Results

Figure 10 is an example of individual performance including less than perfect performance beginning at the number 3 Event Rate Interval block. On average, within that same 2 minute block, fewer than two errors persisted across the block. Within the Figure 10 example, the number 3 Event Rate Interval block represents a desirable level of task difficulty in terms of *er* (events occurring between a min time of 2.94 s and max time of 5.88 s).

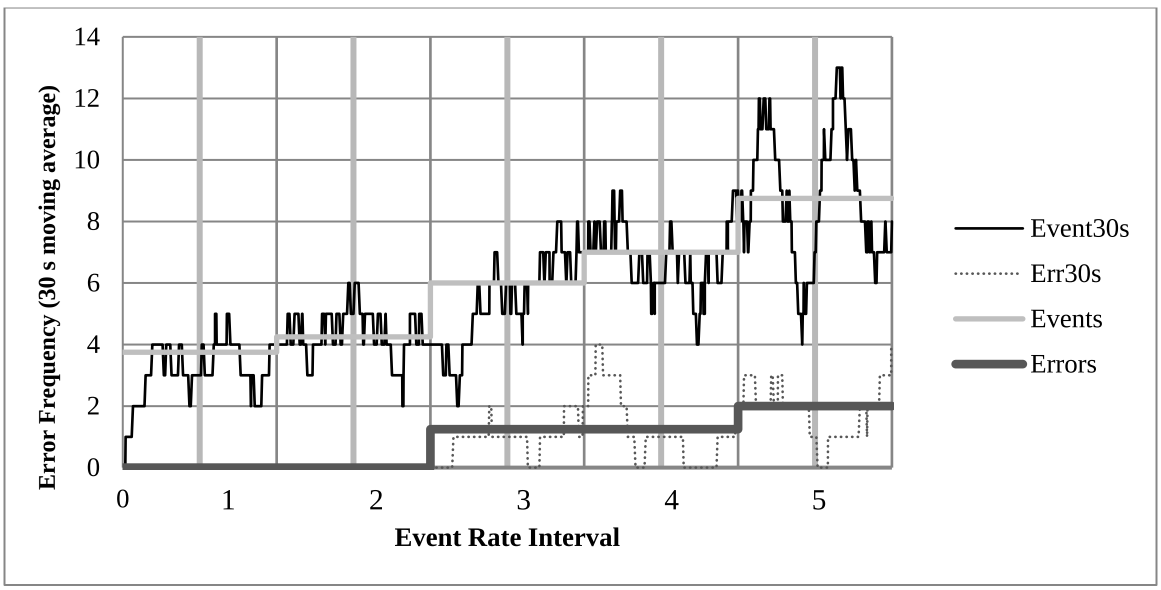


Figure 10. Plotted individual 30 s moving mean performance for a 10 minute trial.

Mauchly’s test indicated that the error data violated the assumption of sphericity for most effects in the experiment, therefore the Greenhouse-Geisser correction was applied to the ANOVA degrees of freedom. The three-factor repeated-measures ANOVA for error rate indicated that the factors of Number of Entities (F (1.4, 12.7) = 30.8, MSE = 12.7, $p \leq .000$, $\varepsilon^2 = 0.77$), Event Rate (F (1.8, 15.8) = 4 3.4, MSE = 23.8, $p \leq .000$, $\varepsilon^2 = .83$), and the interaction of entity and event rate (F (3.2, 11.8) = 8.6, MSE = 6.2, $p \leq .000$, $\varepsilon^2 = .49$) were the only significant factors. Fisher's Least Significant Difference post-hoc tests conducted on the main effects indicated that error rate increased significantly with each increase in the number of entities and event rate. Figure 11 depicts the error frequency means (30 s moving average) for each of the Entity and Event Rate Interval combinations. As this figure shows, error frequency generally increases as a function of both the number of entities and the event rate increase. The only exception is that the error frequency does not increase substantially between event rate intervals of 2 and 3 when only 3 or 4 entities are present.

Note: See Appendix G for additional analysis results.

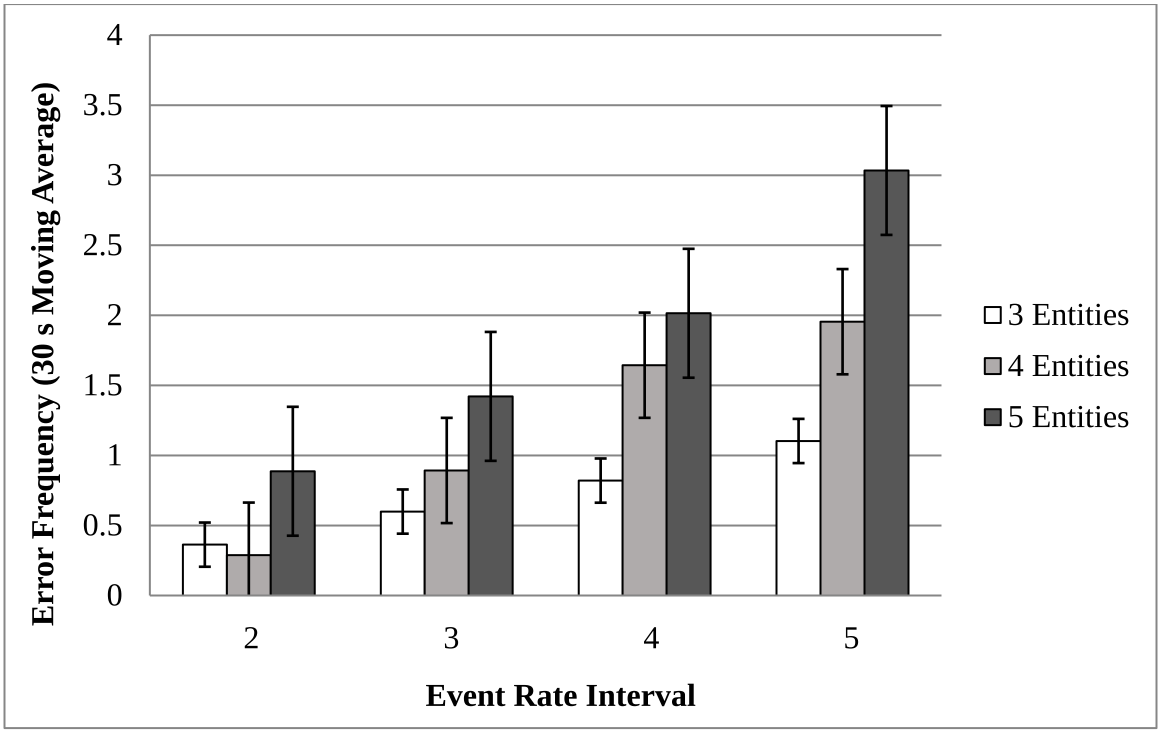


Figure 11. Mean error frequency within collapsed 2 minute Event Rate Intervals across all participants.

A three-way, repeated-measures ANOVA with Greenhouse-Geisser correction conducted on response time indicated that the number of entities, the event rate, the direction, nor any interaction of these variables had a significant effect on response time.

Finally, a two-way, repeated-measures ANOVA with Greenhouse-Geisser correction indicated that the effects of number of entities (F (1.7, 15.2) = 53.6, MSE = 1.7, $p \leq .000$, $\varepsilon^2 = .86$), and direction (F (1.9, 17.4) = 11.4, MSE = 0.15, $p \leq .001$, $\varepsilon^2 = .56$) significantly affected workload. Fisher's Least Significant Difference tests on the main effect of entity indicated that workload increased with each increase in the number of entities. As shown in Figure 12, workload increased steadily from just under 4 (insufficient spare capacity for easy attention to additional tasks) for three entities to nearly 7 (very little spare capacity, but maintenance of effort in the task is not in question) for five entities. A rating of approximately 6 (little spare capacity: level of effort allows little attention to additional tasks) was obtained for four entities. Therefore, it is reasonable to believe that the addition of a task to the current environment with four entities would likely result in performance degradation, making the current task sensitive to the addition of a primary task when the current task is employed as a secondary task.

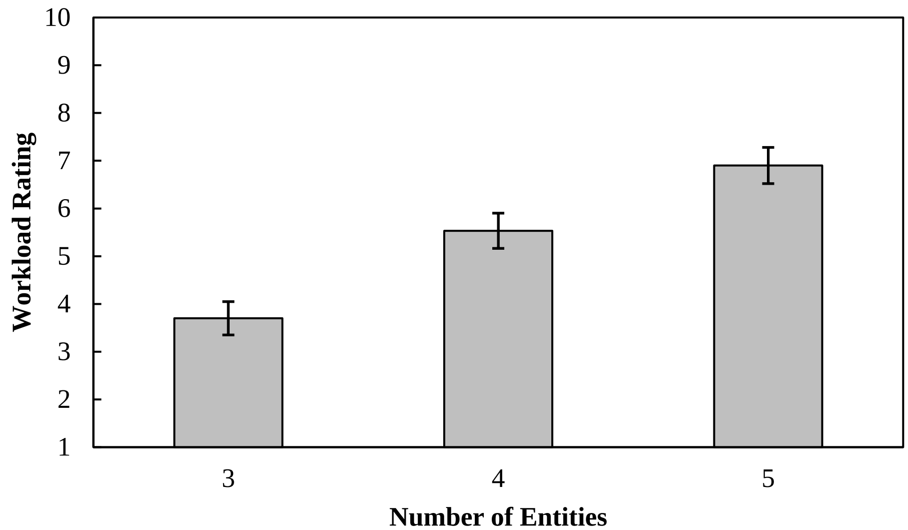


Figure 12. Mean workload rating as a function of the number of entities. Error bars indicate one standard error of the mean.

Fisher's Least Significant Difference tests indicated that workload was significantly higher for direction 3 than directions 1 or 2. However, as shown in Figure 13, workload was relatively flat across direction, increasing from a rating of 5.2 to 5.6 across the directions.

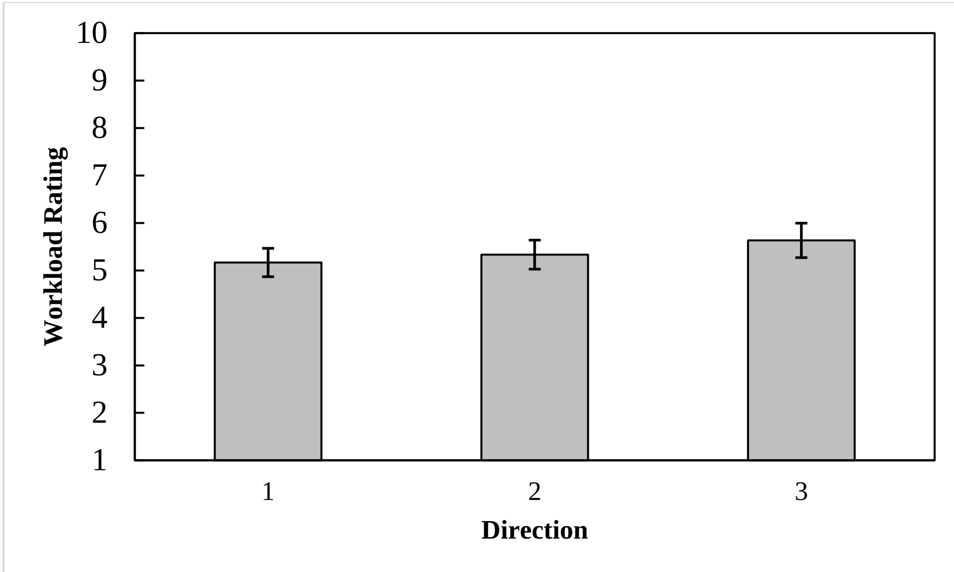


Figure 13. Mean workload rating as a function of entity direction. Error bars indicate one standard error of the mean.

4 Discussion

Application of the dual-task paradigm methodology is well-understood and widely used in human performance experimentation. The secondary task as a performance measurement tool must meet specific criteria to be useful (Fisk et al., 1979; Kramer et al., 1985). The intent of the present paper is to describe the development, mechanisation, and validation of a secondary task set that adheres to the necessary criteria and can be adopted for use in a wide variety of human performance evaluations for visual information portrayal interface comparison evaluations.

Future evaluation activities should include quantitative comparisons between the methodology discussed here and other secondary tasks which have been used in the past. Examples are the well-known Psychomotor Vigilance Task (Drummond, Bischoff-

Grethe, Dinges, Ayalon, Mednick & Maloy, 2005) and the MAT-B (Comstock & Arnegard, 1992).

The described secondary task requires constant visual attention and provides results that are sensitive to lapses in visual attention. Additionally, the results indicate that the task difficulty can be tuneable over a large range by changing the event rate or the number of entities. Further, with careful selection of randomisation seeds, multiple conditions can be provided with nearly equivalent task load. The resulting task can be completed on common desktop computers, could be trained in under 30 minutes, and does not require specialised skills.

The objective performance and subjective workload indicate that selection of four entities and an event rate between a minimum of 2.94 s and a maximum of 5.88 s interval is appropriate for the secondary task difficulty. This condition provides near error-free performance and an average workload rating which indicates that participants have little spare capacity for additional tasks. Therefore, the addition of a primary task is likely to degrade secondary task performance, inducing additional errors. As such, this task is expected to be sensitive to potential performance differences among a manipulation of independent variables within a primary task. This level of performance is consistent with the objective of avoiding a ceiling effect. The fact that participants were successful in correctly reporting changes for a large number of entities is consistent with the objective of avoiding floor effects. Further, selection of four entities is consistent with previous literature which indicate the ability of individual's abilities to visually track up to four entities simultaneously (Pylyshyn & Storm 1988; Oksama & Hyönä 2004). Ideally the adaptation of the methodology presented here enables researchers to avoid the time and effort required to develop custom secondary task sets for their unique experimental activities.

As the proposed standardized secondary task set appears to meet the requirements stated in this paper and within the literature (Fisk et al., 1979; Kramer et al., 1985) for a secondary visual psychomotor task, the authors have future plans to utilise the secondary task set to compare various visual interface concepts within a primary task. This activity will act to further exercise the potential utility of the proposed ‘standardised’ secondary task methodology.

5 Acknowledgements

The authors are sincerely grateful for the participants who volunteered their time and effort toward this experiment.

6 Declaration of Interest

The authors are aware of no potential conflict of interest.

7 References

Castelhano, M. S., & Muter, P. (2001). Optimizing the reading of electronic text using rapid serial visual presentation. *Behavior & Information technology*, *20*(4), 237-247.

Chase, B., Irwin-Chase, H., & Sonico, J. (2004). A methodology to control for individual differences in dual-task performance. *Theoretical Issues in Ergonomics Science*, *5*(5), 445-451.

CIE. (2002). *The correlation of models for vision and visual performance* (No. CIE 145:2002).

Comstock, Jr., J. R., & Arnegard, R. J. (1992). *The multi-attribute task battery for human operator workload and strategic behavior research* (No. NASA TM-104174). *NASA Technical Memorandum*. Hampton, VA.

Drummond, S. P., Bischoff-Grethe, A., Dinges, D. F., Ayalon, L., Mednick, S. C., & Meloy, M. J. (2005). The neural basis of the psychomotor vigilance task. *Sleep, 28*(9), 1059-1068.

Fisk, A. D., Schneider, W., & Derrick, W. L. (1979). A methodological assessment and evaluation of dual-task paradigms. *Psychological Research & Reviews, 5*(4), 315-327.

Kahneman, D. (1973). *Attention and effort*. Prentice-Hall, Englewood Cliffs, NJ.

Kantowitz, B. H., & Sorkin, R. D. (1983). *Human factors: understanding people-system relationships*. Wiley, New York, NY.

Kramer, A. F., Wickens, C. D., Donchin, E., & Fregly, A. (1985). Processing of stimulus properties: evidence for dual-task integrality. *Human Perception and Performance, 11*(4), 393–408.

Oksama, L., & Hyona, J. (2004). Is multiple object tracking carried out automatically by an early vision mechanism independent of higher - order cognition ? An individual difference approach. *Visual Cognition, 11*(5), 631–671.

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets : Evidence for a parallel tracking mechanism. *Spatial Vision, 3*(3), 179–197.

Roscoe, A. H., & Ellis, G. A. (1990). *A subjective rating scale for assessing pilot workload in flight : A decade of practical use* (No. TR 90019). Farnborough, Hampshire.

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors, 50*(3), 449-455.

IV. Experiment 2

Chapter Overview

This study employed a secondary-task paradigm intended to empirically evaluate different methods for accessing text-based information during the performance of an entity monitoring and interaction task. The three different text access methods (RSVP, TS, and CC) were evaluated against a baseline traditional pop-up dialog box (DB) presentation format. The basis of the text information content was a “9-line” standardized targeting information readout discussed in more detail below. The experimental text access methods or, techniques/formats, were designed to provide an operator access to the 9-line information without expanding the display area much beyond the size of what otherwise may be that of an entity identification tag. This was done by presentation of a dynamically “running” display of the 9-line content associated with the appropriate entity via either the RSVP, TS, or CC techniques. Additionally, participants had the option to interact with the text presentation enabling direct control of start/stop, selection of initial start point, as well as movement forward or backward (reversion) through the text. This control over the text content was similar to the use of a miniature multifunction display. As suggested by the literature, a progress meter feature was included with all formats except the baseline.

The participant was required to accurately and quickly confirm entity level information compared to a master inquiry of the 9-line information content. This was called the “query probe” or “query” task. Accuracy and response time were among the dependent measures used as primary task performance metrics. Participants were instructed to prioritize completion of the primary task over that of the secondary task as well as prioritization of accuracy over completion duration in general.

The secondary task was intended to mimic the monitoring and reporting tasks performed by command and control display operators, as described in Experiment 1. Participants reported changes of presence, identification, and direction of movement for multiple entities within the display area. Accuracy and response time were used as performance metrics for the primary and secondary task. Since participants were asked to prioritize the primary task over the secondary task, accuracy and response time differences were hypothesized to not exist for the primary task across the independent variables of interest. Therefore, performance and accuracy in the secondary task across independent variable levels were posited to serve as the objective basis of comparison.

The results from Experiment 1 were used to help determine the "level of difficulty" utilized for the Experiment 2 secondary task. The objective performance, subjective workload, and previous literature (Pylyshyn & Storm, 1988; Oksama & Hyönä, 2004) indicated that selection of the n = 4 entity level and 241s min - 360 s max *er* setting was appropriate for a secondary task difficulty to be sensitive to potential performance differences among the text access format manipulation for the primary task. The difficulty determination objective was to avoid either a ceiling or floor effect among the secondary task variables.

Method

Apparatus

A standard high-resolution (2560 horizontal x 1500 vertical) LCD monitor was utilized to present the visual stimulus to the participants. The active area of the display measured 26.67 cm (1066 pixels) horizontal (h) by 19.87 cm (791 pixels) vertical (v). The active area expressed as visual angle was 23.72 deg h by 17.78 deg v. Participants were seated at a normal visual distance (approximately 60 cm) from the display source. This was representative of the viewing distance for a conventional workstation and the associated level of comfort expected of a seated system operator.

A conventional computer equipped with standard keyboard and cursor input device (mouse) was used to generate the stimulus display and for participant response input. Questionnaire responses were also collected via pen and paper using the questionnaire shown in Appendix B.

Software development for this study was performed in-house by Eric Heft under direction of the author. Software development included generation of all the visual elements, task mechanization, performance data collection routines, raw data recording, and any automatic data reduction.

Participants

Twenty four (24) individuals participated in the experiment. The number of participants was determined by comparison to similar studies where adequate statistical power was demonstrated (Broadbent & Broadbent, 1987). Additionally, the following equation (Brown and Prescott, 2006) was used to determine the adequacy of the planned number participants for this research:

$$n = 2(z_{1-\alpha/2} + z_{\beta})^2 \sigma^2[1+ (m - 1) \rho]/ m \Delta^2$$

where m = number of repeated measures, σ^2 is the between-participant variation, ρ is the correlation between observations on the same participant, and Δ is the difference in error counts to be detected between treatments. Data from the difficulty determination study (Experiment 1) was used to estimate the variance and correlation expected to be observed. Assuming $\alpha = .05$, power of at least 0.80 (i.e., $\beta = .2$), estimating from a first replication where $\sigma^2 = 28.67778$, and $\rho = .4108$ from the average correlation between observations on the same participants, a sample of size 24 participants per treatment was determined to be sufficient to detect error count differences between treatments of at least 3.25 errors; a difference that is considered meaningful for the envisioned application.

Participants were recruited among local area personnel and from a locally managed professional participant pool. The final participant pool included individuals who were active duty, civilian, and contractor employees on Wright Patterson AFB, Ohio. Participant recruitment was completed via “word-of-mouth” and email. No specific skillset was required. Participants were required to self-report normal or corrected-to-normal visual acuity and color vision for inclusion (Appendix A).

Demographic Questionnaire

A questionnaire, consisting of background information questions (e.g. age, vision correction, handedness, computer use, and video-gaming experience) was administered prior to data collection trials (as shown in Appendix A). Participants input their responses using pen and paper. Mean participant age was 26.5 years with a range of 18 to 50 years. Median participant age was 24 years. The participant group included 14 males and 10 females. All participants reported normal or corrected to normal visual acuity and normal color perception. Left handedness was reported by four participants and one reported no hand preference. Reported weekly video game use ranged from 0 to 40 hours. Mean video game use was reported as 9.24 hours per week. Weekly television viewing ranged from 0 to 20 hours with an average duration of 8.17 hours per week. Median television viewing was 10 hours per week. All but two participants reported computer use on a weekly basis for work and all but 5 reported weekly computer use for leisure. Computer use ranged from 1 hour to 80 hours per week. The average reported weekly computer use was 30.81 hours. The median reported computer use was 25 hours per week. Computer use for gaming purposes was reported by 14 of the 24 participants.

Duration

The experimental sequence including reviewing an introduction, informed consent, instructions, completion of demographics and vision screening questionnaire, data-collection, and rest breaks. The total experimental duration was less than 320 minutes per individual. Sessions were completed over multiple data collection days to avoid excessive fatigue during participation.

Text Access Technique Development Description

This section introduces the various techniques and user interaction concepts designed to provide user access to a relatively large amount of readable text in a small display area. The display area of interest was constrained roughly to that which one would expect to be used for an entity label within a C^3I data presentation. Additional display real estate was consumed by the control mechanization graphics added to non-baseline designs. The text content of interest was representative of the information within a standard 9-line format used to pass targeting information between command functions and mission performers.

Example Nine (9)-Line Used in the Context of Close Air Support (CAS)

The 9-line communication protocol objective is to support a targeting or sensor data collection activity. A top-level requirement for this communication sequence is the identification of the mission element acting agent or performer. For C^3I display purposes, this item is typically a call sign label. The 9-line information following the call sign allows command and control functions to understand the assignment and intent of the acting agent (Air Land Sea Application Center, 1997). The 9-line content can be thought of as the critical information items “briefing” required to perform a mission assignment. For consistency and ease of use, the items contained in the 9-line are formatted in a standard sequence. In support of decision making activities, access to the 9-line affords communication at the single entity level. Also, as mission changes occur and new assignments are passed to entities, 9-line information forms the basis of a closed communication loop between command functions and actors. The elements of the standard 9-line sequence are provided in Table 3.

Table 3. Contents and descriptions of 9-line elements.

Identification (ID) tag: call sign label
1) Initial Point (IP) or Battle Position (BP): Ingress point expressed as a landmark, waypoint, latitude/longitude coordinates (lat/long), etc.
2) Heading: From IP/PB to target expressed in magnetic compass coordinates.
3) Distance: From IP/BP to target expressed in nautical miles (from IP) or meters (from BP).
4) Target Elevation: Expressed in feet/mean sea level (MSL).
5) Target Description: As appropriate.
6) Target Location: Visual, lat/long, grid, or offsets.
7) Type Mark: Laser, waypoint, infrared, or beacon. Code: Actual code.
8) Location of Friendly Forces: As appropriate. Position Marked By: As appropriate.
9) Egress: As appropriate. Remarks: As appropriate. Time On Target: Expressed in Universal Time Constant (UTC).
Remarks: Additional information as needed.

Baseline Display Condition (Pop-Up Dialog Box (DB) Technique)

For purposes of comparison, a baseline text access technique, or format, consisted of a conventional pop-up dialog box approach. Each active entity in the aggregate display had an associated ID data tag attached to it. The data tag box was a fixed size which was able to accommodate approximately 13 characters. The tag box (ID tag) size was based on the human ability to perceive 12-14 characters in a single fixation with acceptable accuracy (Just & Carpenter, 1998; Robeck & Wallace, 1990; Rayner, 1998; Öquist, 2006). Figure 14 (as presented by Öquist, 2006) demonstrates the foveal and parafoveal accuracy approximation within a single human eye fixation (perceptual span).

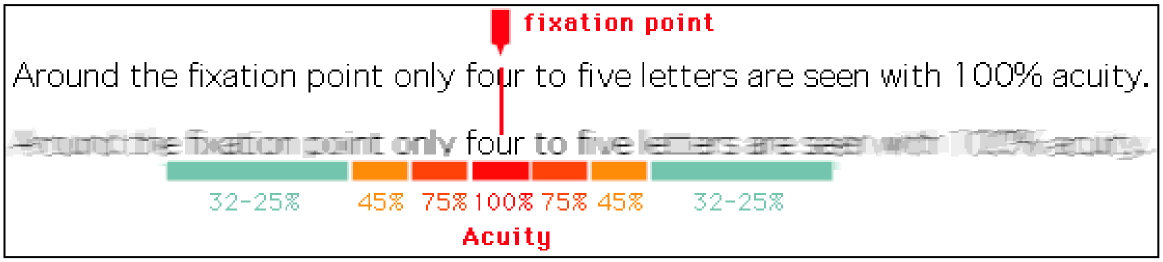


Figure 14. Single fixation foveal and parafoveal perceptual accuracy (Öquist, 2006).

For this investigation, it was desirable that the entity label be perceived with a single fixation and that the alphanumeric characters of the ID tag require little to no saccadic eye movement to be read in their entirety (Figure 15). For access to the complete 9-line data associated with the entity, participants could expand the ID tag by selection with a conventional pointing device (hover the mouse-driven cursor over the ID tag and left-click). This action resulted in the display of a pop-up dialog box containing the entire 9-line information set for that entity. The vertical size of the 9-line dialog box varied depending on the amount of text required by the 9-line assignment. If necessary, text was wrapped to fit within a window size defined by the often used "page" format 50 character-wide limitation (Piolat, Roussey, & Thunin, 1997). Figure 16 shows a rendering of a pop-up dialog box. The window remained open until another left-click within the window was registered by the mouse controller. This interaction feature enabled the user to keep multiple windows open for each of the multiple entities as long as desired. It was expected that this interaction mechanization was familiar and intuitive for users and thus required little or no usability training.

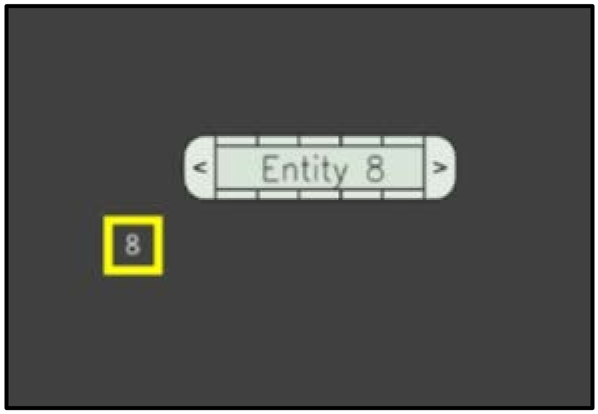


Figure 15. Entity with identification data tag label.

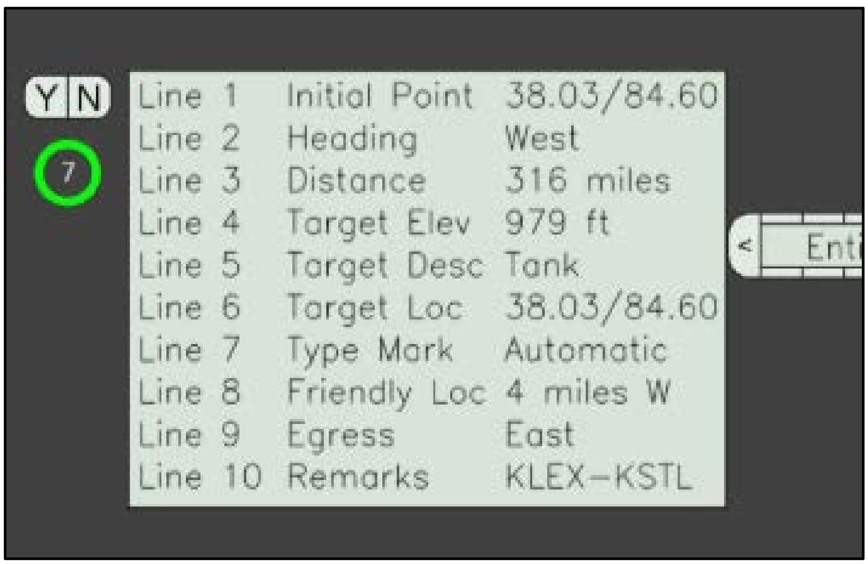


Figure 16. Example entity with pop-up dialog box containing 9-line information.

Spatially-Constrained Concepts

Figure 17 shows a rendering of a spatially-constrained entity tag including user interaction symbology. Similar to the baseline condition, the drawing area of the data tag accommodated approximately 13 characters at a time. This was the maximum size of the drawing area regardless of the mechanization of the underlying different text presentation types. It was assumed that some ID label must remain with each entity simultaneously with the cycled 9-line information. This was be accomplished by superimposing an ID number within each entity shape (Figure 18).

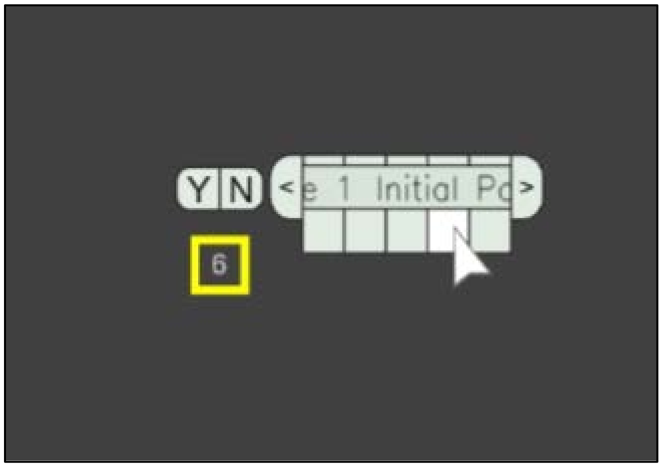


Figure 17. Example spatially-constrained presentation window and function "soft keys".

Figure 18. Entity shape with the addition of an ID number.

When selected by the participant, the 9-line information associated with the entity of interest became active and the presentation continued until the entity was "de-selected" by the user. This functionality allowed the information associated with multiple entities to be available simultaneously. Text presentation techniques considered included: DB, RSVP, TS, and CC (Table 4). For purposes of evaluation, the DB format represented the non-interactive baseline text access technique.

Table 4. Text presentation techniques included in the evaluation.

Text Presentation Variables			
Pop-up Dialog Box (DB)	Rapid Serial Visual Presentation (RSVP)	Times Square (TS)	Cinema Credits (CC)

Interaction and Feedback Concepts

To enable information seeking tasks, i.e., those requiring the location of information from specific lines, interaction elements were designed to support the spatially-constrained text access formats. Accordingly, some of the most important empirical evaluation questions were related to the performance trade-offs which exist as a result of this interaction. For instance, it was anticipated that the DB mechanization may pose significant occlusion cost, but information access was likely to be quite quick. Artifacts of this presentation approach may make it difficult to maintain situation awareness of multiple entities compared to the spatially-constrained (SC) variants. The following are some of the functions that were supported by the interface (Figure 17):

Direct Selection

The upper and lower bezel of the SC window was segmented into 10 small “soft keys” which were selectable by the participants. Hovering the mouse cursor over the soft keys resulted in their graphical expansion so that selection was potentially easier and more accurate than with the original size of the keys. Again, this design was conceptually similar to interaction with a multifunction display. Each of the bezel keys corresponded with a respective 9-line element (Table 3). The 10th key was reserved for any “remarks” information that the entity 9-line included. When the participant selected one of the keys on the bezel (via a mouse left-click), the corresponding 9-line text was presented dynamically via the appropriate access technique being evaluated.

Run Function

This function was activated by left-clicking within the area of the SC presentation window. When activated, the 9-line text associated with the entity cycled within the drawing area via one of the presentation technique formats. Going from the static label state to the run state, the 9-line text began with the first category line and proceeded in sequence to the end of the 9-line information. This cycle was repeated until some other mode was selected by the participant. Cycle duration was set at 20 s from beginning to end. This translated to a presentation rate of approximately 15 characters per second.

Pause / Stop Function

When the presentation was actively running, a left-click within the drawing area "froze" the dynamic presentation of text and displayed a static string of text that was visible within the dialog box at the instant the left-click selection was made. If the user took no other action for 5 seconds (s), the mode switched automatically to a "stop" state and only the top level entity label text was presented within the SC presentation window.

Fast Forward Function

The rightmost vertical bezel component provided a fast forward soft key. This function skipped the text presentation to the beginning of the next line in the 9-line sequence when pressed during the run or direct selection mode. Depending on the initial mode prior to this action, the respective presentation continued from the new start point.

To activate a fast forward, the user hovered the mouse-driven cursor over the fast forward soft key and left clicked the mouse input button.

Reversion Function

The left vertical bezel provided a reversion soft key. A reversion action during run mode restarted the 9-line sequence again from the beginning of the present text line if selected toward the last 2/3 of that line. If selected during the beginning of the text line (first 1/3), the reversion action snapped to the previous line and began to run. To activate a reversion, the user hovered the mouse-driven cursor over the reversion key and left clicked the mouse input key.

Feedback or Completion Meter

A shade fill technique was be used to provide the user simultaneous orientation and progress information. Within each of the bezel keys a left to right fill animation indicated which line was being presented and, the amount of fill indicated progress within the text string of the current line. For instance, a half-filled key indicated half of that line of text has been presented and half was yet to be presented. Line presentation completion corresponded to the key being completely filled. When the presentation switched to the next line, the fill animation began again for that line key. This feature was intended to provide a means for participants to become quickly oriented to location within the 9-line text. Also, this approach was intended to provide a smooth and unobtrusive cycle progress tracking capability.

Presentation Window Entity Dependence and Interaction

When the text presentation window of a given entity was selected, the window remained at a fixed location in screen coordinates, possibly permitting easier interaction, or the window continued to move with the entity. This behavior was treated as a separate independent variable with two levels. One level acted as if the text presentation window was “attached” to the entity. This level was called dynamic dependence (DD). The second level fixed the x and y coordinates of the text presentation window once selected. This level was called static independence (SI).

Dynamic Dependence (DD) Interaction Mechanization

The term “dynamic” indicates that the text access window was free to move within the x and y coordinates of the boundaries of the scenario area including exit from the scenario area. The term “dependence” indicates that the text access window moved in conjunction with its parent entity in such a way that it was tethered to the upper right side of the entity as shown in Figure 17. Text access interaction (see the Interaction and Feedback Concepts section in this chapter) was controlled by the participant via a combination of hovering the mouse cursor over the entity data tag and associated mini multifunction display soft keys. Activation of the interaction control features was accomplished by a left clicking the left button on the mouse. Multiple entity text access data tags could be activated at any time by moving the mouse cursor among any of the entities present.

Static Independence (SI) Interaction Mechanization

The term “static” indicates that the text access window became locked to a specific x and y coordinate within the boundaries of the scenario area when activated by the participant. The term “independence” indicates that, once activated, the text access window was independent of the continuous movement of its associated entity. The entity was free to exit the scenario area while any activated text access window remained active.

Activation of the interaction control features was accomplished by clicking the left button on the mouse. Multiple entity text access data tags could be activated at any time by moving the mouse cursor among the present entities. A fast double click (i.e., having a second click within 750 ms) of the left mouse button within the text access window acted to “snap” the data tag back to its originally associated entity. This included removal from the scenario area if the original associated entity had exited the scenario area while the text access window was active.

Occlusion and Layering

Relative to each other, all entities, data tags, and text access windows were opaque. To assign occlusion, a layering approach was used based on the order in which entity interaction was sequenced. The most recent interaction brought those objects (set combinations of entities and their associated data tags and activated text access windows) to the top layer and sent each other object one layer back from its previous layering sequence. Similarly, entity presence order determined layer position in the absence of user interaction. The most recent object with the most recent interaction remained at the top or “front” layer. The occlusion sequence described above was the same regardless of whether the text access interaction mechanization was DD or SI.

Experimental Design

The study was a 4 by 2 by 4 full-factorial within-subjects experimental design. There were four levels of text access technique (DB, RSVP, TS, and CC) and two level of interaction mechanization (DD and SI) (as depicted in Table 5). The third variable, initial direction, was included to add variation and a feel of randomness to the secondary task. This variable was intended to make it difficult for participants to recognize any pattern to the presentation of the secondary task entities even though the number of events and their timing within the secondary task was similar across all trials. The four levels of the initial direction variable were formed by beginning a trial so that the first entity entered from the original side border of the gaming area frame or from one of the three remaining sides. This flipped or reversed the direction of the sequencing of all of the subsequent entity behavior for the duration of that trial. For the full-factorial design, each of the 24 participants participated in 32, 600 s data collection trials.

Order effect confounding for the repeated measures variables was mitigated by fully counterbalancing using a Latin Square design, as described by Keppel (1982). The General Linear Model, repeated measures, Analysis of Variance (ANOVA) tests were completed using SAS JMP and IBM SPSS. These ANOVAs were performed separately for the response time and accuracy dependent measures. Post hoc analysis using Tukey honest significant difference (HSD) tests were performed for comparisons where statistically reliable main effects or interactions were indicated. Given the interdependency between response accuracy and response time measures, an analysis of the independent variables using the response duration measures was performed while holding the accuracy variable constant so that accuracy was always 100% correct for that measure.

Table 5. Independent variables matrix.

Text Access Format				Interaction Mechanization
DB	RSVP	TS	CC	DD
DB	RSVP	TS	CC	SI

Secondary-Task Paradigm for Performance Measurement

Kantowitz and Sorkin (1983) gave a very simple but logical explanation for human information processing mechanism upon which the secondary task performance measurement methodology was based. Central to this concept is that humans have a limited capacity to process information and perform tasks. This is particularly true when tasks require similar modalities or processing channels for their performance (Wickens, 2008). When there is a requirement that two tasks be performed simultaneously, humans can perform the tasks without decrement provided the “pool” of resources available for the effort is not exceeded by the difficulty of the effort (Kahneman, 1973). If the effort to perform two tasks simultaneously exceeds the available resources, some performance decrement across the tasks is expected. If one of the tasks is prioritized by the human as primary, it follows that any performance decrement present will be isolated to the secondary task. As a methodology, this well-established phenomenon can be used to measure any differential capacity requirements of the primary task which may be attributable to an experimental manipulation. For this research, given no systematic differences other than a manipulation of the text access techniques (DB, RSVP, TS, and CC using DD or SI), any significant performance decrement on a secondary task can be interpreted as a result of different levels of “usability” existing among the levels of the text access independent variables.

The present research used the secondary task methodology to detect any measurable performance differences among the text access techniques and entity dependence levels by recording participant performance on a secondary task that was common across all combinations of the independent variables of interest, shown earlier in Table 5. By instruction and training, participants were required to treat the primary task, where they used the text access techniques, as their performance priority. They were instructed to perform the secondary task to the best of their ability without sacrificing performance on the primary task.

The following paragraphs describe the primary and secondary tasks and their associated performance measures. The tasks were designed to mimic those which are required of some operational command and control tasks (e.g., air traffic control radar, battle management information display, flight deck tactical situation displays, etc.).

Primary Task

The primary task was designed to mimic an operational workstation where it is necessary for the operator to confirm specific information contained within the 9-line data associated with an identified entity of interest. Metaphorically, the task was designed to be representative of a simple query where a superior communicates an information request to a system operator. In real-world operations, a communication like this could take many forms and is likely to be auditory or verbal in nature. For this research, it was desirable that the modalities involved for task completion (i.e., input and output interaction) were as similar to one another as possible and highly controlled. For these reasons, the information query task took the form of a text-based probe. The query contained the identification of the entity of interest, the 9-line category where the information was contained, and the specific data level information to be confirmed. Figure 19 is an illustration of how a text probe query was presented to the participants (Entity 7 Line 5 Target Desc Launch Pad).

When an information query was presented to participants, their task was to access the appropriate information for the identified entity of interest and confirm that the information associated with the entity was an accurate match by moving the mouse over and selecting the “Yes (Y)” input or that the information associated with the entity is not an accurate match by moving mouse over and selecting the “No (N)” input. A small input box was presented so the Yes/No input could be recorded. Two buttons within the box were labeled with a “Y” for reporting a Yes response and an “N” for reporting a No response. When the cursor was hovered over one of the Y/N buttons, the button nearest the cursor tip was highlighted. Whatever button was highlighted when a mouse button was clicked was recorded and logged as an associated input. Again, Figure 19 includes an illustration of the confirmation input selection Y/N buttons, referred to as the query report buttons. The query report buttons were present while a valid query was displayed and were removed after an input was made or the query reached the timeout duration. The total number of positive and negative but accurate responses was balanced across the trial blocks while presentation query response type was randomized.

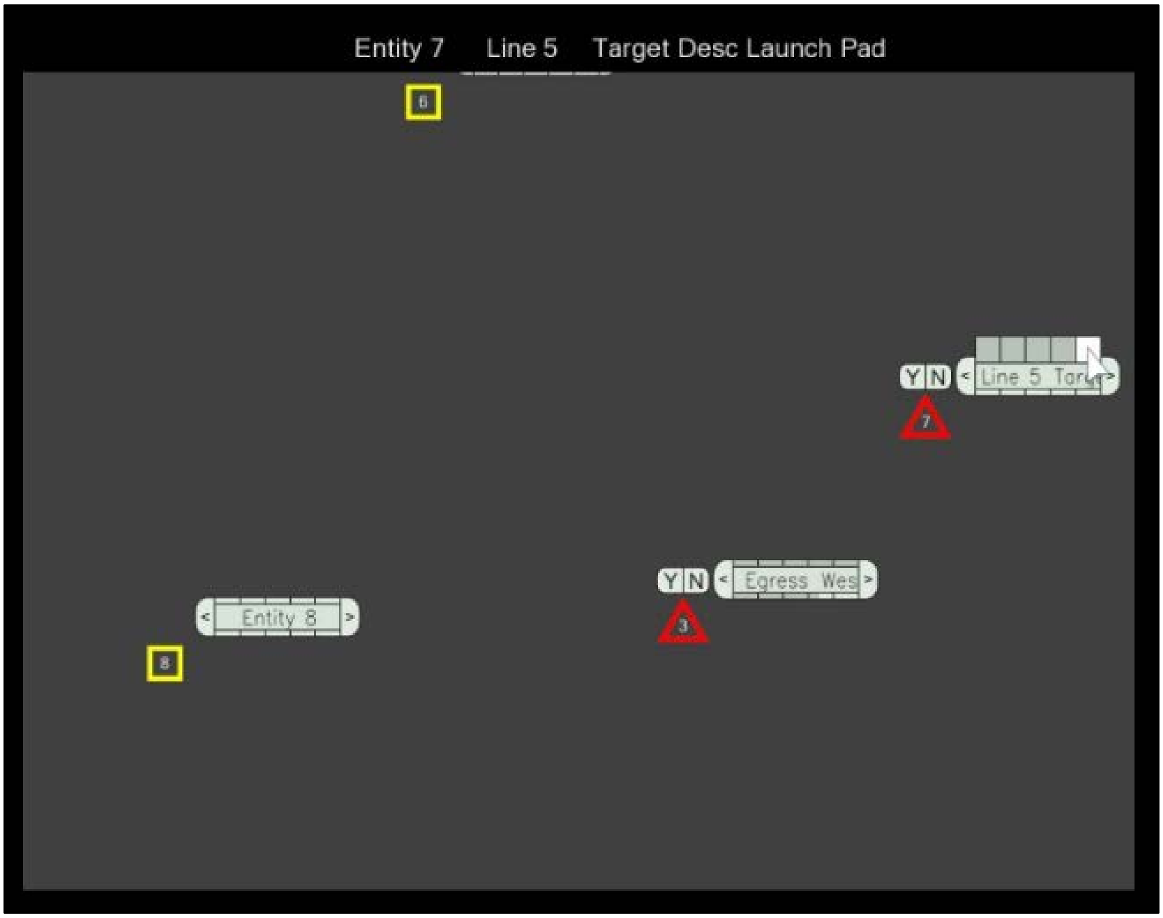


Figure 19. Example primary task query probe.

By explicit instruction, participants were requested to treat the primary task as their highest priority when presented with a query probe. Further, to avoid unwanted tradeoffs in speed and accuracy, the instructions asked the participants to prioritize accuracy over speed when performing the query task. Participants were directed to perform the secondary task only to the extent possible given spare capacity to do so during completion of the primary task.

Nine (9)-Line Information

The 9-line data generated for the study was selected so that it met specific criteria. The data values were simple but realistic and were made to appear to be operationally relevant information as if the entities were individual aircraft with tactical assignments. The data were formed of latitude, longitude, and elevation coordinates, heading directions, distances, target types, locations, etc. The data were selected to be easily and quickly understood as well as easily held within the limitations of working memory by minimally trained naïve participants. No specific skill set was required to perform the tasks. Further, the number of individual text characters within a query probe data value was designed to be accurately perceived by the participant within the limitations of a single foveal fixation. This was based on the normal human ability to perceive 12-14 characters in a single fixation with acceptable accuracy (Robeck & Wallace, 1990; Just & Carpenter, 1998; Rayner, 1998; Öquist, 2006).

The data shown in Table 6 was the basis for both the query input and the 9-line content associated with the specific entities displayed in the four different presentation techniques and two different levels of the dependency interaction mechanization. For a positive or “Yes” correct confirmation, the query and entity 9-line data items matched. For a negative or “No” correct confirmation response, the query item represented the matching category of information in terms of line number and description but the selection entity and the query entity were assigned different values. Participants were not able to determine a correct query response by recognizing that the query data item was

Table 6. Nine (9)-line data content.

Line 1 Initial Point	**Line 2** Heading	**Line 3** Distance	**Line 4** Target Elev	**Line 5** Target Desc	**Line 6** Target Loc	**Line 7** Type Mark	**Line 8** Friendly Loc	**Line 9** Egress	**Line 10** Remarks
38.37/81.59	North	163nm	1202ft	Runway	40.49/80.23	Visual	8nm NW	South	KCRW-KPIT
42.40/83.01	Northeast	521nm	626ft	Fuel Tank	42.40/83.01	Laser	5nm NE	Southwest	KDET-KBUF
39.04/84.66	East	413nm	896ft	Hangar	39.04/84.66	Infrared	8nm SW	West	KCVG-KRIC
36.19/95.88	Southeast	230nm	677ft	Landing Pad	36.19/95.88	Beacon	6nm NE	Northwest	KTUL-KLIT
36.08/98.15	South	239nm	2181ft	Bunker	36.08/98.27	Talk on	5nm S	North	KLAS-KNYL
33.63/84.42	Southwest	147nm	1026ft	Tower	33.63/84.42	Radio	7nm S	Northeast	KATL-KMGM
38.03/84.60	West	316nm	979ft	Tank	38.03/84.60	Auto	4nm W	East	KLEX-KSTL
39.90/84.21	Northwest	241nm	1009ft	Launch Pad	39.90/84.21	Code	9nm N	Southeast	KDAY-KORD

the wrong category of information for the query line number or description. A correct response included confirmation of matching data items when the query entity and the selected entity were also the same. Similarly, a correct response was recorded when the participant selected "No" when the query entity were the same but the data values did not match.

Query Event Control

Query events occurred according to a controlled schedule. Each 10 minute (i.e., 600 s) data collection session included a total of 22 queries, which equated to one query every 25 s. The first 25 s period of a trial did not include a query to ensure enough time was allotted to allow the scenario to be fully populated with secondary task entities. Additionally, a "blank period" of 25 s occurred at the 300 s point of the trial to make the query timing pattern less predictable. Finally, the initialization of each query was randomized within the first 5 s of each query period. The remaining 20 s of each period was allotted for completion of the query task. If the task was not completed within that 20 s window, the query task timed out and was recorded as such.

Within the 22 queries, half (11) of the correct responses were data matches ("Yes" responses) and half (11) were data mismatches ("No" responses). The selection of entity related to a probe was spread approximately equally across the entities within the scenario (22 probes per trial / 4 entities in the scenario = approximately 5.5 queries per entity, per trial). The selection of the query entity was randomized without replacement with the following further logical constraints: 1) the selected entity was present at the time of query and 2) the entity was not scheduled to exit the scenario within the following 10 seconds.

Primary Task Systems Modeling Language (SysML) Activity Diagram

Figure 20 depicts the activity elements encountered during performance of each instance of the primary task. The initialization point in the diagram indicates that the secondary task was interrupted by a query probe which appeared according to the schedule discussed in the previous section. The first step of the activity was selection of the correct entity specified within the query by clicking the left mouse button after hovering the mouse cursor over the desired entity. Until the correct entity was selected, the activity remained within a loop at this step. The following step was to access the correct line which contained the value to be compared to the queried value. Depending on the text access steps chosen by the participant, clicking the left mouse button on the proper multifunction soft key resulted in a dynamic jump to the associated text line. Now, the participant was able to compare the entity text line value with the query value to make a Yes (match) or No (mismatch) input. These selections were made with a left or right mouse click. After this input was completed, or 20 s elapsed, that instance of the primary task was terminated.

After primary task termination, the diagram indicates continued secondary task performance. The secondary task events persisted during performance of the primary task and participants could chose to shift between the primary and secondary tasks at will. By instruction and training, participants were encouraged to prioritize performance of the primary task when a query probe was active. This included abandoning performance of the secondary task while a primary task was underway. The ability to shift between the tasks is depicted in Figure 22 where the secondary task is shown as an interruptible activity.

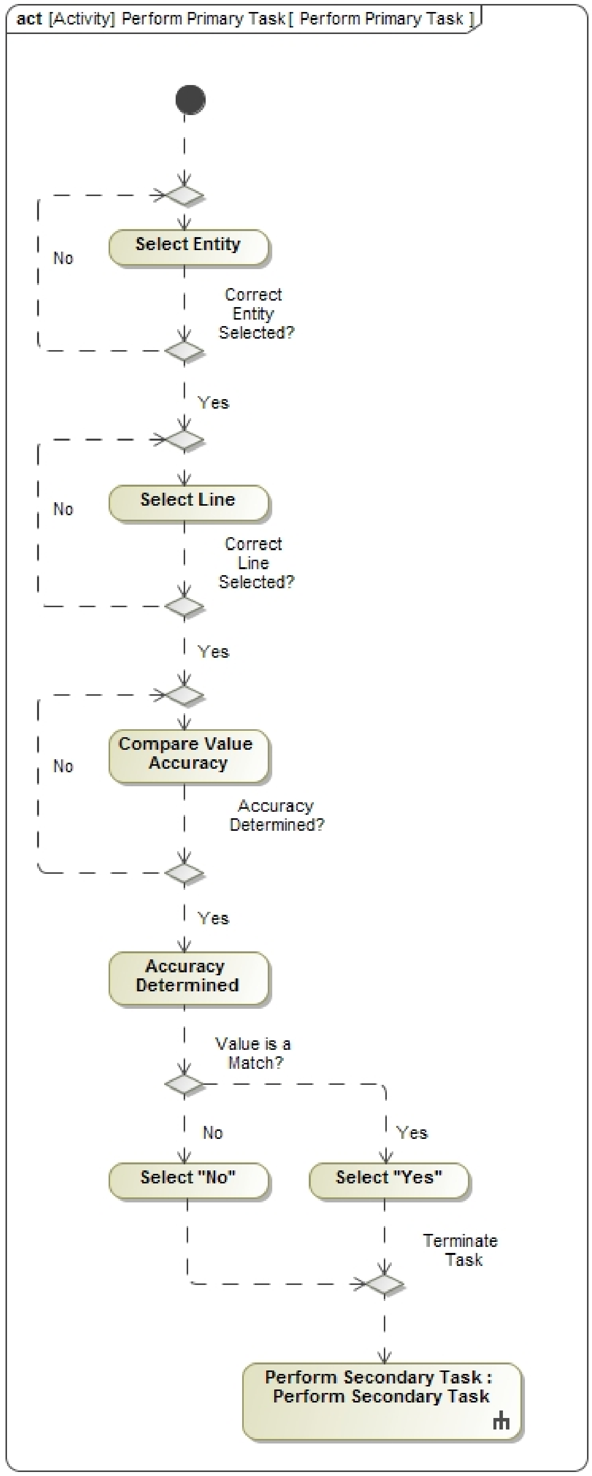


Figure 20. Primary Task SysML Activity Diagram.

Primary Task Dependent Measures

The primary task dependent measures of interest for this research included response accuracy and response duration (time). The accuracy measure for the primary task included correct responses (i.e. correctly confirming that the query and target data values matched when they were supposed to and the query and target did not match when they were not supposed to), or incorrect responses (indicating that the data values matched when they in fact did not, or indicating the data values did not match when in fact they did). These responses were recorded as frequency of correct and incorrect responses and compared as a proportion of correct responses.

When a query probe was presented, the system clock count was initiated and the following duration categories were recorded based on cursor location and mouse click: time to entity text tag selection (correct or otherwise), time of text access interaction initiation, and time to task completion. These measures can be combined to give the elapsed response time of any combination of the intervals.

The following primary task dependent measures were recorded for purposes of analysis and findings interpretation:

1) Accuracy measures: correct, incorrect, and incorrect type.

2) Response time measures: total response time per query task and elapsed time from query presentation to entity selection.

3) Timeout occurrences.

4) Subjective Workload: recoded at the session or trial level (text access format with entity dependence).

5) Text access technique subjective preference feedback was recorded after the participant had experienced all of text access techniques (Appendices B and C).

Secondary Task

The secondary task was designed to mimic a generic scenario monitoring task performed by operators such as air traffic controllers or battle management analysts. The task required participants to monitor the activity of several entities and report all specified observed entity status changes when they occurred. Figure 21 is a screen capture showing the presence of three entities in the scenario. The entities moved across the screen at a constant rate in one of four possible directions relative to screen coordinates: up (+Y), down (-Y), right (+X), or left (-X). Entity identification was differentiated by standard shape and color coded symbols: green circle, red triangle, and yellow square. Each type included its associated entity ID number within the center area of the entity shape and a data tag was attached to each symbol (Figure 21). The data tag was the component where the primary task text access occurred.

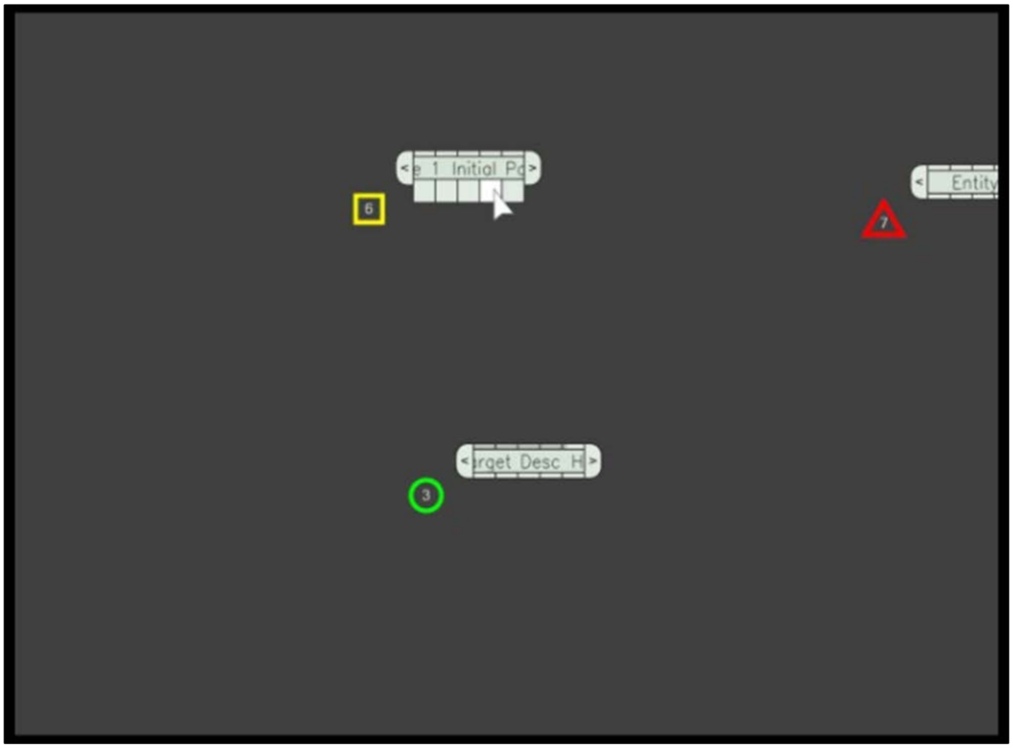


Figure 21. Multiple entity identification symbology, data tags, and cursor.

Participants were instructed to monitor multiple entities for three specific events that were “reportable” within the scenario. Entrance of entities into the scenario from “off screen” was reportable as “presence” when they became visible. Once the entities were present, participants monitored the entities for both changes in direction and changes in identification. At controlled intervals, entities became present and, once present, the motion of an entity shifted from its current trajectory to a +/-90 degree or -180 degree change from the current course. Any time a direction change occurred, it was a reportable “direction” change event. Finally, a change in entity “identification” was reportable. When this change occurred, the entity changed shape and color to become one of the other two types of entities. The identification number within the entity shape did not change. Participants were instructed to report any time a shape identification change was detected. Table 7 shows the three different reportable entity events within the secondary task activities. The following paragraphs describe how the secondary task events were scheduled and how participants interacted with the secondary task.

Table 7. Reportable entity events for the secondary task.

Entity Change Type		
Presence (P)	Identification (I)	Direction (D)

Secondary Task Event Control

Entities entered and exited the scenario according to a controlled schedule. No change events occurred during the first 15 s of runtime to allow the scenario to be populated with entities. After that, the scheduling logic sequenced through all entities that were not currently in an active event status, were not scheduled to exit the scenario within 10 s given their current trajectory, or had not had the same event occur within the preceding 10 s. If no entities met the event criteria, a “blank” event was recorded. For an otherwise eligible entity, an event was assigned according to the following probabilities: direction change (p = .40), identification change (p = .40), and no event assignment (p = .20). The frequency of events was controlled by a min/max event timer so it was controllable with some precision. Similarly, the scenario was designed to maintain a fairly constant mean number of active entities at any given time once the general population was established at about the first 15 s point. A logical schedule was employed here as well. If the scenario was underpopulated by one, a new entity was added within a 10 to 20 s interval. If the scenario is underpopulated by two or more and the next entity was not scheduled to enter for more than 10 s, then an entity was scheduled to enter within a 5 and 10 s interval. If the scenario logic detected that two entities would exit the scenario within the next 10 s, and no new entities were scheduled to enter within the next 10 s, a new entity was scheduled to enter the scenario within an interval of 2 and 4 s.

Secondary Task SysML Activity Diagram

Figure 22 depicts the activity elements to be performed by the participant during the secondary task. The initialization point indicates a 600 s trial is underway. In general, the secondary task required the participant to monitor the entities and report on their presence in the scenario as well as changes of identification and direction. The change events occurred according to the schedule discussed in the previous section. Participants selected entities and reported change events by hovering the mouse cursor over the desired entity and actuating a right or left mouse click to input a response. The diagram indicates that the secondary task activity could be interrupted by a primary task query probe as depicted in Figure 20. Participants were encouraged by direction and training to prioritize performance of the primary task when the interruptions occurred.

The trial was terminated when there were no longer any entity events to be reported, i.e., there was no active query activity, and the 600 s endpoint time limit had been reached.

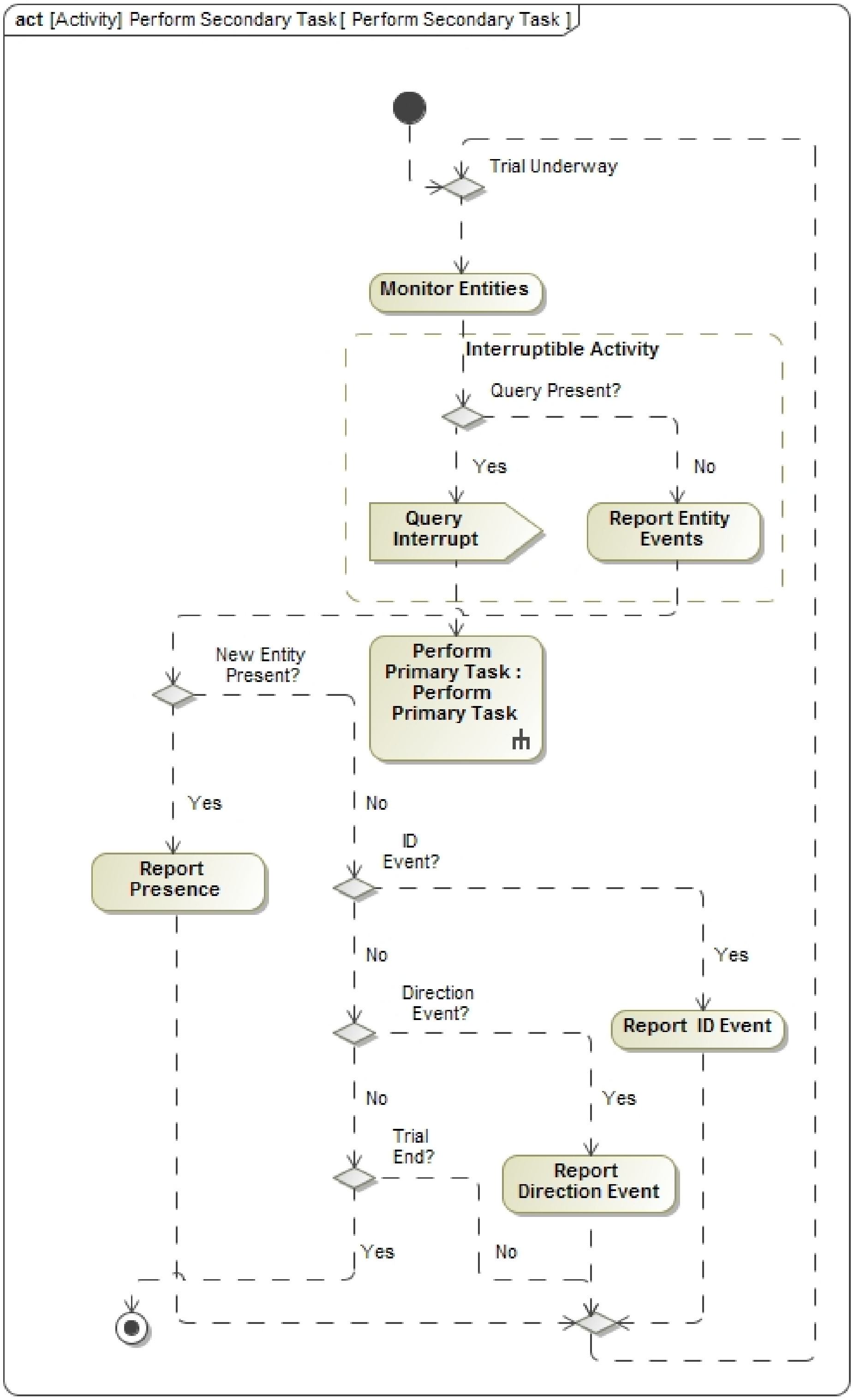


Figure 22. Secondary Task SysML Activity Diagram.

Secondary Task Difficulty

The intent of the event schedule and its associated monitoring functions was to generate the appearance of random events for participants. In fact, the events were scheduled and controlled. The desire to control event scheduling was two-fold. First, the basis for a constant set of scenario events could be presented to participants in a way that they were not likely to predict the sequence of events. This was accomplished by making event initialization dependent on participant interaction and the ability to initialize a common scenario with different initial entry direction and location. Second, difficulty of the secondary task could be manipulated directly by controlling the average number of entities in the scenario and the rate of change assigned to the event schedule probabilities within the bounds of the control criteria described above. Of course difficulty could be manipulated by speeding up or slowing down the entity rate of motion as well. As indicated in the previous chapter, it was desirable to tune the difficulty of the secondary task so that, when performed in isolation, it was engaging but near 100% accuracy performance was achievable. This formed the basis of the secondary task measure which was intended to be sensitive to the spare capacity afforded to participants given the simultaneous performance of the primary task and the manipulation of the independent variables within the primary task.

New Entity Initial Condition

When a new entity was added, the simulation scheduling logic assigned a random direction. However, the direction was constrained so that a new entity would not enter from the same side of the scenario gaming area as the previous entity. This was designed to encourage the entities to be spread out. The scheduler then assigned a random position for the entity along the border where it would enter.

The monitor did not check to see if any existing entities would overlap. This potentially made it difficult for participants to click a new entity but the layering scheme was used so that the new entry was "on top" of the otherwise interfering entity. Similarly, any selected entity was moved to the top layer of the scenario. Due to the speed of the entities and their size, an entity entering from the left or right was not completely visible before the 10 s event hold expired but presence reporting could still be accomplished by selecting any visible part of an entity. The following section gives more detailed information regarding the actual event reporting process.

Participant and Task Interaction

All direct interaction between participants and the scenario was accomplished with a standard mouse input device. The mouse was used to move a cursor to any location within the display area within the bounds of the scenario gaming area defined by the outside border depicted in Figure 21. Proximity of the cursor to an entity automatically associated that entity to the cursor for purposes of selection and interaction. A "right" click of the mouse button was used to select the entity for change reporting and an initial left mouse button was used for primary task interaction. Once an entity was selected by a right click, a small input box was presented at the tip of the cursor arrow so the entity change could be recorded. The three buttons within the box were labeled with a "P" for reporting presence of a new entity, an "I" to report an identification change, and a "D" for direction change reporting. Once an entity was selected and the reporting buttons were present, the box remained stationary relative to screen coordinates, which were fixed in x, y screen coordinates position, until a right or left click was input with the cursor hovered over the P, I, or D button within the "PID box". When the cursor was hovered over one of the PID buttons, the button nearest the cursor tip was highlighted. Whatever button was highlighted when a mouse button was clicked was recorded and logged as an associated input. Figure 23 depicts the static PID box which appeared because “Entity 8” was selected with a right click. In the screen capture, Entity 8 has continued to move away from the location where it was selected. The location of the entity selection was approximately where the PID box is drawn. The PID box remained visible until a P, I, or D button was selected by a mouse click (PID choice condition), a right or left click was made near another entity (re-selection condition), a right or left click was made away from any entity or PID box (cancelation condition), or 10 seconds elapsed without any input made by the participant (timeout condition). Only one PID box was present at any one time during a trial.

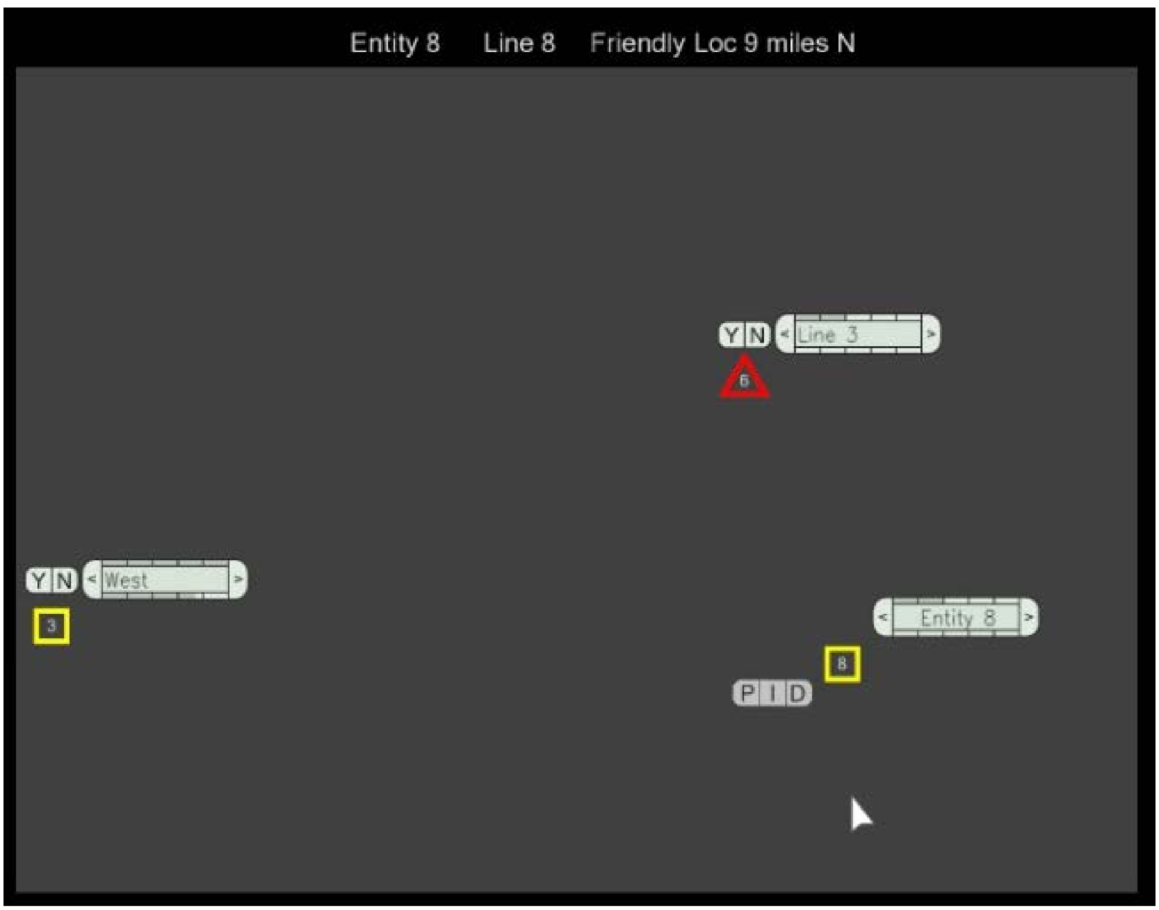


Figure 23. Multiple entities and a “PID box” waiting for participant input.

Secondary Task Dependent Measures

The secondary task dependent measures were also focused on accuracy and response time. The response accuracy measures for the secondary task were based on proper selection by the participant using mouse click input using the PID box. An accurate input was defined by the selection of the proper entity and PID box button for the preceding change event. Errors were defined several ways: 1) failure to select an entity within the maximum allowable time of a PID event; 2) failure to select the correct PID event change button within the maximum allowable time for the correct entity; 3) selection of the proper entity but selection of the incorrect PID event button; or 4) selection of an entity not associated with a preceding change. The proportion of correct responses was used as the operative accuracy measure.

Similar to the primary task response duration measurement, once a change event occurred for a specific entity, a system clock was initiated for that entity. Response time was categorized as time to correct entity selection, time to a response, correct reply response time, or when timeout was reached. Note that incorrect entity selection was registered as a type of accuracy error and the response time for incorrect selections were precluded from the response time calculation. Timeouts, which were also a type of accuracy error, were recorded if the maximum allowable time for registration of a correct response elapsed before a selection was made.

The following secondary task dependent measures were recorded for purposes of analysis and findings interpretation:

1) Accuracy measures: correct entity selection, incorrect entity selection, correct PID selection, incorrect PID selection, and change event misses.

2) Response time measures: total response time per change event, elapsed time from event change to entity selection, elapsed time from entity selection to PID box appearance, and elapsed time from PID box appearance to its removal.

3) Timeout occurrences.

Other Measures

There were other response measures which were of interest to ensure a thorough, interpretable, and informative analysis. For instance, cursor activity is an objective measure which is likely to co-vary with the other measures already discussed and it may add to the ability to make inferences regarding the utility of the different display formats. Cursor activity as a continuous measure may inform how much work the participant had to perform to achieve a level of accuracy within some response duration. This measure lends itself to investigation in terms of movement area, location relative to other objects, and rate of input. At the end of active cursor input associated with a change or query event (indicated by its motion and subsequent stop), the location of the cursor on the display screen may be a good approximation of participant instantaneous gaze fixation. This also may be informative in terms of participant task engagement and overall effort.

In general, all events were recorded with enough temporal resolution to allow the data collection sessions to be replayed at system clock update rate along with all system input actions. This enabled the data reduction step to be revisited in the case where new and unanticipated measures were desired for analysis.

Task Training and Feedback

After completing a written and verbal description of the experimental tasks, participants were trained on the primary, secondary, and combined tasks until the effects of learning on performance appeared asymptotic. Participants were trained first on the secondary task procedures until performance was nearly error free and workload was acceptable. Participants were instructed to maintain a continuous visual scan pattern so that missing change events among the entities was minimized. Next, participants were familiarized with the primary task by performing it in isolation of the secondary task. Participants experienced the primary task with each of the independent variable levels. Once participants were comfortable with the primary task and performance appeared stable, the primary and secondary tasks were combined to train simultaneous performance of the tasks. It was reiterated to participants that they prioritize performance on the primary task and, within both tasks, prioritize accuracy over speed. Participants were reminded to perform the secondary task to the extent possible given any spare capacity to do so.

During the training sessions, participants were provided feedback to help inform them firsthand of their accuracy on the primary, secondary, and combined tasks. Training performance and feedback were monitored by the experimenter so that errors could be pointed out and tips for optimal performance of the tasks could be communicated and standardized across participants. The objective for all participants was to perform every task as similarly as practical so that their performance was representative of trained operators. This process acted to ensure that all participants had the motor dexterity and skill to perform the experimental tasks. No individuals were rejected from participation due to the inability to perform the tasks acceptably.

Subjective Feedback Data Collection

Additional subjective measures were recorded. It was important to collect subjective workload measures at appropriate points within the overall test matrix to record a global measure of participants' perceived level of effort based on their perceived spare capacity across primary and secondary task activity. Subjective preference data was also collected to analyze consistency with the other measures. It is not uncommon to find user interface comparisons where objective measures are inconclusive or inconsistent with otherwise strong participant indications of preference.

Workload Assessment

Subjective workload was recorded by having participants complete a Bedford Workload Scale (Roscoe & Ellis, 1990) decision tree procedure. The Bedford self-assessment workload rating scale is a fairly self-explanatory procedure where participants report subjective task workload based on consideration of the spare capacity to perform additional tasks with regard to the performed tasks. Use of the Bedford Scale was described during participant introduction and instructions. A practice data collection session included demonstrating the use of the scale. Figure 24 shows the Bedford Workload Scale decision tree.

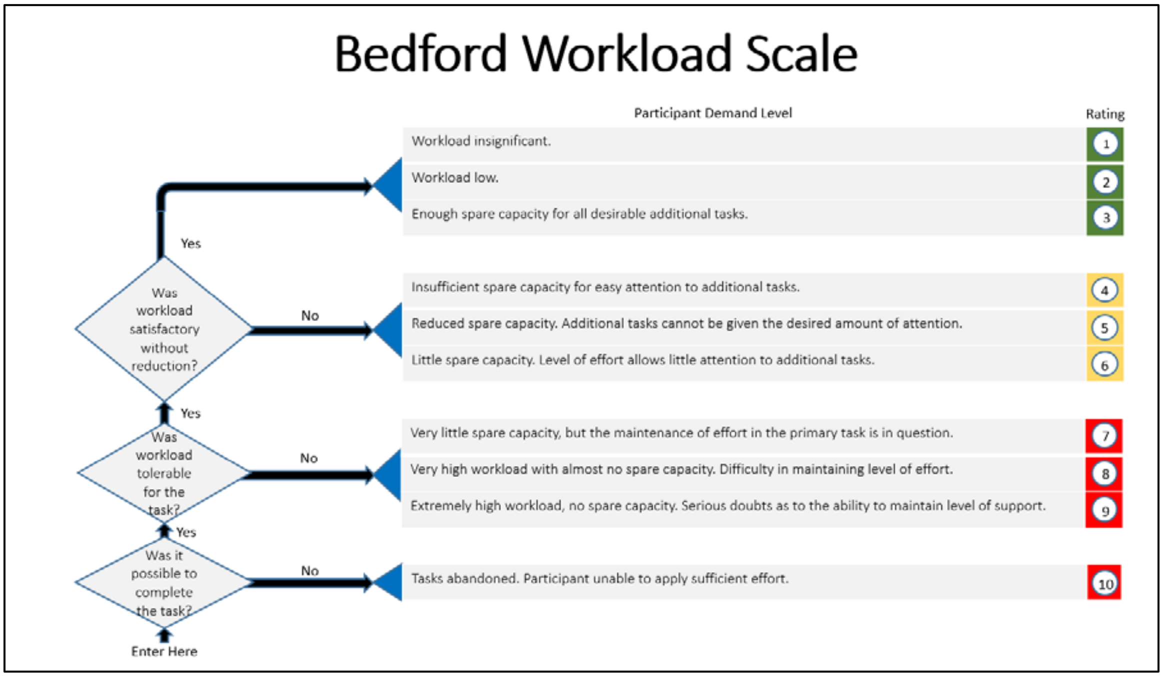


Figure 24. Bedford Workload Scale graphical decision tree.

Subjective Preference Questionnaire

After completion of all data collection sessions, participants completed a rating questionnaire designed to record their subjective preferences for the combinations of the independent variables related to the performance of the primary task. Additionally, participants were asked to indicate how well they thought they were able to perform the primary task using the combinations of text access formats and interaction mechanizations. This questionnaire is shown in Appendix B.

Experimental Control to Reduce the Effects of Individual Differences

A tendency to trade speed for accuracy is an individual characteristic which could vary among the participants, adding unwanted systematic differences or error to the response variables within this study. This variability could reduce the ability to isolate effects of the independent variables.

Additionally, there were other potential environmental influences on task performance which were minimized to the extent possible to reduce systematic error. This was accomplished to preserve generalizability while allowing the task set to be representative of the defining attributes of real-world applications.

Experimental control for individual variability can take the form of participant selection criteria based on appropriate combinations of physical capability (visual acuity, handedness, etc.), skillset (past experience), training, and/or demonstrated performance (motor dexterity and skill). Where these variables were not controlled directly or where unwanted systematic variably was suspected but unavoidable, data regarding these differences was recorded for later analysis. Potential influential conditions which were less controllable may have existed as well. Examples are participant motivation, pre-existing fatigue, different task completion strategies, and differing priorities. These influences could have varied within individuals across the course of participation in the study.

For this study, the intent was that no unusual or specialized skillset be required for participation. Few physical requirements beyond normal or corrected-to-normal visual acuity and reported normal color vision were required. Because the tasks were performed using a conventional workstation display and control interactions, there was little need to control for handedness but these data were recorded via the demographics reporting procedure. There was no reason to suspect gender or age differences but these data were recorded nonetheless. Similarly, a basic notion of past experience was collected through demographics reporting of experience levels and daily exposure to computer workstation type tasks and video game experience. In terms of motivation, it was not anticipated that any performance-based reward would be necessary. Performance feedback provided during training was used to help motivate participants to continue to work toward performance improvement.

In general, acceptable motivation was assumed due to the volunteer nature of participation. Mitigation of strategy and performance prioritization variability (response accuracy versus response speed) was attempted through instruction and training. For example, while conducting Experiment 1, it was found that it was important to emphasize an entity-by-entity scan pattern strategy instead of utilizing a centralized fixed gaze approach for the task. Again, participants were instructed to prioritize the accuracy of their responses over the speed with which their responses were made. Also, in an attempt to mitigate individual skill differences among participants, as well as ensure that they were adequately familiar with the tasks, training to a baseline level of performance, as observed by the experimenter, occurred prior to formal data collection. The utilization of a within-subject experimental design was another individual difference mitigation technique. Counterbalancing was used to avoid order and learning effects. Fatigue and its influence on attention, motivation, etc. was mitigated by including proper rest and recovery periods between data collection sessions. A 10 minute session duration was used to avoid unacceptable sustained effort.

Differences within the data collection environment were controlled, including lighting, noise level, temperature, desk/chair ergonomics, and interface with computer input devices. Data collection occurred in the same physical location using the same lighting conditions and workstation, including display and input devices. Participant posture was allowed to vary somewhat to represent how each individual’s comfort preference would be established naturally for the use of similar input devices for similar tasks. The workstation display location was fixed on the desktop surface.

Data Collection Procedure and Sequence

Participants were seated at a table equipped with a desktop monitor, keyboard, and mouse. The intent was to mimic the environment depicted in Figure 1. The experiment consisted of the following sequence: 1) introduction/consent, 2) demographic questionnaire, 3) vison screening 4) task training, 5) data collection, 6) workload assessment, and 7) preference questionnaire.

1) Introduction/safety briefing/consent: Participants read a short description of the study rationale and sequence of events. They then received a short safety briefing explaining what to do in the event of an emergency. This briefing included exit location identification, tornado, and shelter-in-place locations, and where to assemble if the building was evacuated. Participants read the consent form and signed it to confirm their agreement to participate in the study. Participants were informed that their participation in the study was completely voluntary but dependent on their reporting normal or corrected-to-normal vision. Participants were informed that they may choose at any time to terminate their participation.

2) Demographic questionnaire: Participants completed a questionnaire to report past experience with video displays and tasks similar to those used in the study.

3) Vision screening: Participants were asked to report normal or corrected-to-normal visual acuity and normal color vision. Corrective lens were used for all phases of participation in the study if worn normally for similar computer tasks.

4) Task training: participants were given detailed instructions regarding the task operation, objectives, priorities, and desired performance. Participants were given practice during fully-dynamic task sessions. The practice sessions included specific instruction for, and hands-on experience with, all interactions they performed during the actual experimental events. The entire range of input and output

variations were covered during this training period. Training typically lasted fewer than 40 minutes. During the training session, participants were encouraged to stop and ask questions concerning the tasks. During this time, participants practiced completing the Bedford Workload Scale self-assessment procedure (Figure 24). Once the experimenter and participant were satisfied that the tasks were fully understood and adequate performance had been demonstrated, the data collection session were initiated.

5) Data collection: Each participant completed 32 data collection sessions (4 text access formats by 2 interaction mechanizations and replications (4)) which each lasted 10 minutes (320 minutes of actual data collection time). The tasks consisted of the actions described above within the Primary and Secondary Task sections.

6) Workload assessment: Workload reporting was performed by participants completing a Bedford Workload Scale decision tree procedure, as depicted in Figure 24, at the completion of each of the discrete 10 minute sessions.

6) Preference Questionnaire: Participants completed the preference questionnaire, which was designed to elicit subjective preference regarding the levels of the independent variables included in the study. Once complete, participants finished their participation in the study.

Results

Overview

Experiment 2 was conducted as a 4 by 2 by 4 full-factorial within-subjects (repeated measures) experimental design. There were 4 levels of text access technique (DB, RSVP, TS, and CC) and two levels of interaction mechanization (DD and SI) independent variables of interest (Table 5). The final 4 levels were the initial direction the secondary task entities entered into the scenario (left, right, top, or bottom). Initial direction remained in all of the analyses to ensure that it as a factor did not interact systematically with any other manipulated independent variable.

General Analysis Strategy

The collected data were reduced and formatted appropriately for statistical analysis. Descriptive measures were output for purposes of general discovery and data exploration. The exploratory data analysis step included the following: examination of the dataset for missing data points, lack of timeline synchronization, and multivariate correlations. Distribution plots were output to identify potential problems, such as dataset outliers, lack of distribution normality with transformations applied as required, and variance sphericity using Mauchly's test of sphericity. When indicated, Greenhouse-Geisser degrees of freedom corrections were applied in response to lack of sphericity.

Three factor repeated measures, Analyses of Variance (ANOVA) tests were conducted for statistical hypothesis testing across the independent variables for each of the dependent variables of interest. For purposes of hypotheses testing, alpha (a) was set at $p < .05$ to reject the null hypothesis. Effect size was estimated by calculating the Partial Eta Squared (η_p^2) using the procedure described by Richardson (2011) for determination when the sample size is not small. Post hoc analysis using Tukey honest significant difference (HSD) tests were performed for comparisons where statistically reliable main effects and/or interactions were indicated.

Multivariate Correlations

Spearman's rank correlation coefficients (r_s) and associated comparison probabilities (p) were calculated to explore the relationship among the main dependent variables recorded for both the primary and secondary tasks in Experiment 2. Table 8 shows the correlation matrix and associated comparison probabilities for the accuracy and response time measures. A moderately strong negative correlation was indicated between the primary task response time for correct responses measure and the number secondary task percent correct responses (r_s = -.69). There was also a moderately strong positive correlation between the secondary task response time for correct responses measure and the primary task response time for correct responses (r_s = .56). Lastly, a negative correlation was shown between the secondary task percent correct responses and the secondary task response time for correct responses (R_s = -.54).

Table 8. Primary and secondary task dependent measures correlation matrix.

Dependent measures	Primary Task Percent Correct Responses	Primary Task Response Time for Correct Responses	Secondary Task Percent Correct Responses	Secondary Task Response Time for Correct Responses
Primary Task Percent Correct Responses	r_s = 1.00			
Primary Task Response Time for Correct Responses	r_s = -.11 p = .0023	r_s = 1.00		
Secondary Task Percent Correct Responses	r_s = .14 p < .0001	r_s = -.69 p < .0001	r_s = 1.00	
Secondary Task Response Time for Correct Responses	r_s = -.02 p = .56	r_s = .56 p < .0001	r_s = -.54 p < .0001	r_s = 1.00

Appendix F includes the descriptive statistics and distribution plots for the four major dependent measures within the analyses: primary task mean percent correct responses, primary task mean response duration for correct responses, secondary task mean percent correct responses, and secondary task mean response duration for correct responses. There were no missing data points detected within the completed dataset and timeline synchronization was intact.

Repeated measures factorial General Linear Model ANOVA procedures were completed and examined for statistical significance among main effects and interactions. Due to the apparent violation of normality in some of the distributions, analyses were run with log or square root transformations. In no case was the analysis outcome different due to these transformations. The decision was made to utilize the original data distribution because the normal distribution violation appeared to be minor and the interpretation of the findings would be simplified by avoiding the data transformation step.

Upon inspection of potential data outliers, it was concluded that no data removal was justified. There was no indication that outlier data were erroneous and the nature of the secondary task paradigm and overall difficulty of the experimental task could understandably cause occasional response times which were lengthy.

Inferential Tests

ANOVA F-test procedures for repeated measures were performed separately for the response time, accuracy, and subjective dependent measures across the primary and secondary tasks. To analyze effectiveness given the interdependency between response time and response accuracy, the accuracy variable was held constant such that the effect on response duration was recorded and analyzed only for correct responses.

Primary Task

A three-way within-subjects repeated measures ANOVA was conducted to compare the effect of text access technique, interaction mechanization, and initial direction for the mean percent correct probe response for the primary task. Text access technique had a statistically significant effect on percent correct probe response [F (3, 69) = 3.88 MSE = 34.68 p = .0127 η_p^2 = .144]. Post hoc comparisons indicated that the mean percent correct probe response for the cinema credits technique (M = 83.45%, *SD* = 5.58%) was significantly higher than both RSVP (*M* = 81.94%, SD = 5.93%) and times square (*M* = 81.7%, *SD* = 5.69%) text access techniques. There was not an indicated significant percent correct probe response difference between the pop-up dialog box (*M* = 82.62%, *SD* = 5.54%) technique and any of the other three levels of the variable. Table 9 shows the mean percent correct responses and the comparisons among the text access techniques for the primary task.

Table 9. Primary task text access technique comparisons for percent correct.

Text Access Technique	Mean Percent Correct (%)	Post Hoc Comparison
Cinema Credits	83.45	A
Pop-Up Dialog Box	82.62	AB
RSVP	81.94	B
Times Square	81.7	B

A similar ANOVA procedure was performed for the response time measure for correct probe responses. Text access had a significant effect on response time [F (3, 69) = 15.38 MSE = 0.27 p < .0001 η_p^2 = .401]. The mean response time for the pop-up dialog box technique (*M* = 5.86 s, *SD* = 1.02 s) technique was significantly quicker than any of the other tested techniques. The times square technique resulted in the next quickest mean response time (*M* = 6.36 s, *SD* = 1.28 s). The RSVP technique resulted in the slowest mean response time (*M* = 6.7 s, *SD* = 1.35 s) while the cinema credits technique response time (*M* = 6.48 s, *SD* = 1.26 s) did not differ significantly from either the times square or RSVP techniques. Table 10 shows the mean response time for correct responses and the comparisons among the text access techniques for the primary task.

Table 10. Primary task text access technique comparisons for response time.

Text Access Technique	Mean Response Time for Correct Responses (s)	Post Hoc Comparison
Pop-Up Dialog Box	5.86	A
Times Square	6.36	B
Cinema Credits	6.48	BC
RSVP	6.7	C

A two-way interaction for response time was also indicated between the text access technique and interaction mechanization variables [F (3, 69) = 4.34 MSE = 0.27 p = .0073 η_p^2 = .159]. The post hoc comparison is shown in Figure 25. For the static level of the interaction mechanization variable, there were no statistically significant differences for mean correct probe response time between any of the text access techniques except the pop-up dialog box level (M = 5.94 s, SD = 0.1 s), which was significantly quicker than the response time for the RSVP condition (M = 6.58 s, SD = 1.3 s). Alternatively, at the dynamic interaction level, the pop-up dialog box response time (M = 5.8 s, SD = 1.05 s) was significantly quicker than all of the other text access technique levels. For the same measure, RSVP response time (M = 6.81 s, SD = 1.4 s) was slower than all other levels with the exception of the cinema credits text access technique (M = 6.64 s, SD = 1.31 s).

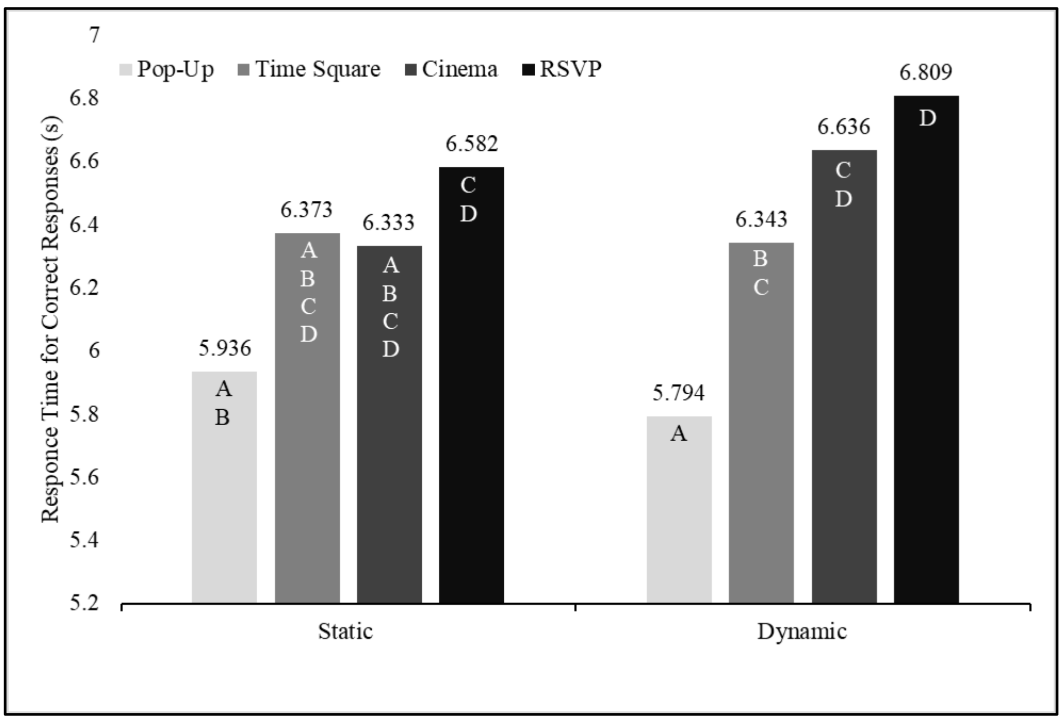


Figure 25. Post hoc comparisons of primary task correct probe mean response time by interaction mechanization two-way interaction.

Secondary Task

A three-way within-subjects ANOVA was conducted to compare the effect of text access technique, interaction mechanization, and initial direction for mean percent correct responses for the PID-based secondary task. There was a significant main effect of text access on percent correct response [F (3, 69) = 4.62 MSE = 17.94 p = .0053 η_p^2 = .167]. Post hoc comparisons indicated that the mean percent correct PID response for the pop-up dialog box text access technique (M = 81.97%, SD = 11.5%) variable was significantly more accurate than all of the other text access technique levels. There were no other indicated differences among the other levels of text access technique levels: RSVP (M = 79.5%, SD = 13.19%), times square (M = 79.6%, SD = 12.86%), or cinema credits (M = 79.37%, SD = 12.68%).

A similar ANOVA procedure was performed for the response time measure for correct PID responses. There was a significant main effect for text access on response time [F (3, 69) = 4.39 MSE = 0.03 p = .0069 η_p^2 = .160]. Post hoc comparisons indicated that the mean response time for the pop-up dialog box technique (M = 2.79 s, SD = 0.41 s) was significantly quicker than the cinema credits text access technique condition (M = 2.88 s, SD = 0.41 s). The times square (M = 2.82 s, SD = 0.40 s) and RSVP (M = 2.86 s, SD = 0.41 s) response time for correct PID response did not differ significantly from either of the remaining two levels.

Because of the fairly strong negative correlations between some of the task measures, which is typical of human speed accuracy tradeoffs, another ANOVA procedure was performed on the secondary task inverse efficiency score (IES) formed by taking the ratio of the response time for correct responses and the corresponding percent correct responses (Vandierendonck, 2018). This test was conducted to detect if any outcome change would result from integrating secondary task accuracy and response time measures into a single measure where a lower IES is indicative of better performance.

There was a significant main effect for text access on IES [F (3, 69) = 4.54 MSE = 0.00002 p = .0058 η_p^2 = .17]. Post hoc comparisons indicated that mean IES for the pop-up dialog box technique (M = 0.035, SD = 0.0094) was significantly lower than the cinema credits (M = 0.038, SD = 0.0106) and RSVP (M = 0.038, SD = 0.0116) text access technique conditions. The times square technique (M = 0.037, SD = 0.0105) did not differ significantly from any of the other levels.

Table 11 shows the comparison differences among the secondary task dependent measures including the IES integrated measure.

Table 11. Secondary task comparisons and dependent measures

	Secondary Task Dependent Measure (Mean)					
Text Access Technique	Percent Correct (%)		Response Time (s)		IES	
Pop-Up Dialog Box	A	(81.97)	A	(2.79)	A	(0.0352)
Times Square	B	(79.6)	AB	(2.82)	AB	(0.037)
RSVP	B	(79.5)	AB	(2.86)	B	(0.0376)
Cinema Credits	B	(79.37)	B	(2.88)	B	(0.0378)

Table 12 shows a recalculation of multivariate Spearman’s rank correlation coefficients and associated comparison probabilities among the primary task dependent measures and the secondary task IES measure. There was a fairly strong positive correlation indicated between the primary task response time measure and the secondary task IES measure.

Table 12. Primary task dependent measures and secondary task IES correlation matrix.

Dependent measures	Primary Task Percent Correct Responses	Primary Task Response Time for Correct Responses	Secondary Task IES
Primary Task Percent Correct Responses	r_s = 1.00		
Primary Task Response Time for Correct Responses	r_s = -.11 *p* = .0023	r_s = 1.00	
Secondary Task IES	r_s = -.10 *p* = .0069	r_s = .70 *p* < .0001	r_s = 1.00

Cursor Activity

A high-level measurement of cursor activity was derived to detect differences in the amount of movement required for the participants to perform the experimental tasks across the different the text access techniques. A variable called cursor distance was calculated by recording cursor movement within model space for each 600 s trial. An ANOVA procedure was performed on these data and found a significant main effect for text access technique [F (3, 69) = 3.02 *MSE* = 24,481.51 p = .036 η_p^2 = .12]. Post hoc comparisons indicated that cursor distance for the pop-up dialog box technique (*M* = 1151.23, *SD* = 389.78) was significantly higher than the RSVP access technique level (*M* = 1084.28, *SD* = 327.88). Times square (*M* = 1123.69, *SD* = 411.78) and cinema credits (*M* = 1093.08, *SD* = 343.54) cursor distance did not differ significantly from the remaining two conditions.

For additional context, the “model space” units were transformed to dimensional screen distance in centimeter units by dividing the mouse distance model space units by the total model space. From this, visual angle was derived using a 63.5 cm notional viewing distance to the screen surface. Dividing the outcome of the formula by trial time gives the average distance measure in cm / s. A similar calculation was performed to transform model space to visual angle and rate expressed as deg / s. Table 13 shows the output of these calculations for each of the text access techniques

Table 13. Cursor distance model space units transformed to real-world units.

Text Access Technique	Average Cursor Distance (cm)	Average Cursor Rate across the screen (cm / s)	Average Cursor Rate in visual angle (deg / s)
RSVP	3450.81	5.66	5.03
Cinema Credits	3478.81	5.70	5.07
Times square	3576.22	5.86	5.21
Pop-Up Dialog Box	3663.89	6.01	5.34

Display Area

Utilizing the frame-by-frame recoding of the experiment trials, a post-processing routine was written to analyze the potential effects of the text access techniques on the area of background occluded by entities and text access graphics over time. One operative question being: "what was the effect of the pop-up dialog box versus the other techniques in terms of display real estate area occupied for the different access techniques during performance of the primary and secondary tasks across the time to complete a trial?". This metric was derived by counting the "background" pixels every 100 ms and calculating non-background pixels drawn and expressing the output as mean percentage of the area occupied by all else other than background across a trial. This percent residual background metric was produced as a way to measure the potential for the text access techniques to occlude points of interest within the total area of the scenario gaming area (background). Table 14 shows the calculation results for the different graphics features which make up the expanded pop-up dialog box (10.66% area), the other three text access techniques (0.92% area), and entity symbol (0.14% area) expressed as the percent residual background measure when static and with only one present in the scenario.

Table 14. Static percent residual background represented by graphics features.

Text Access Technique	Static Percent Residual Background Area (%)
Pop-up dialog box technique	89.34
Other techniques	99.08
Entity symbol	99.86

From Table 14 it can be appreciated that the pop-up dialog box text access technique had an almost 10 fold impact on percent residual background occupied versus the other techniques. It is also the case that that during the data collection trials, the pop-up dialog box was expanded for only a short period of time. To account for the difference in display time in the analysis, percent residual background was output as a mean across trial time (dynamic). Table 15 shows the results of the mean effect of each of the different text access techniques across a 600 s trial.

Table 15. Mean dynamic percent residual background area for each text access technique.

Text Access Technique	Mean Percent Dynamic Residual Background Area (%)
RSVP	96.61
Cinema Credits	96.62
Times square	96.63
Pop-up Dialog Box	95.61

While the difference among the text access techniques is small (approximately one percent), the data indicate it is likely that the pop-up dialog box condition reliably occupies more of gaming area than the three remaining text access techniques even when expansion time is considered. These data were examined and it was discovered that distribution of residuals was heavily skewed and the overall shape of a fitted normal curve was also quite flat. Reasonable transformations were not successful as a remedy for the normality assumption so it was decided to abandon overall residual background as a meaningful measure of area effects in favor of a measure to capture the interaction of the text access techniques with the secondary task entities themselves.

Secondary Task Entity Occlusion

Again, the pixel image-level analysis was used to derive an output measure of entity occlusion defined as frequency count of instances where a direction or identification entity change was "mostly" occluded by overlaying graphics of the primary task text access technique graphics. In this case, "mostly" was defined as 75% to 99% occlusion of the secondary task entity by text access graphics. A two-way, within-subjects ANOVA was conducted to compare the effect of text access technique and interaction mechanization on the number of secondary task entity occlusion cases across the period of a trial and found a significant main effect for text access technique [F (3, 69) = 92.37 MSE = 0.559 p < .0001 η_p^2 = .80]. Post hoc comparisons indicated that primary task occlusion by the pop-up dialog box technique (M = 0.67, SD = 0.6) was significantly higher than the other techniques: times square (M = 0.323, SD = 0.469), cinema credits (M = 0.323, SD = 0.48), or RSVP (M = 0.328, SD = 0.471). These last three techniques did not differ significantly from one another. Figure 26 illustrates the comparison of the pop-up dialog box technique versus the other techniques for 75% to 99% entity occlusion occurrences during direction or identification change events.

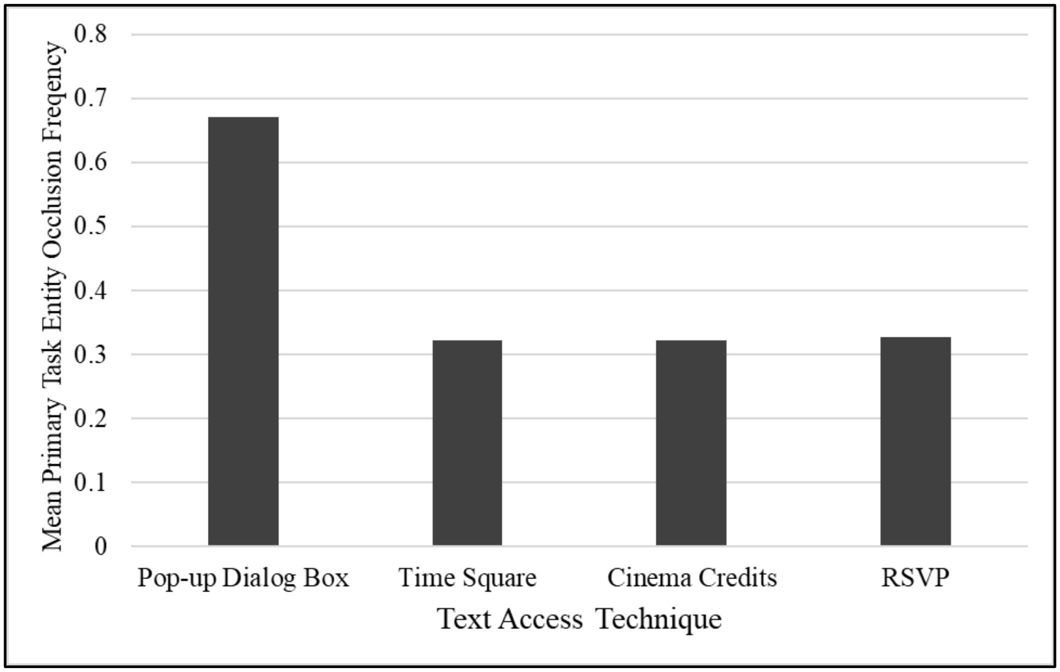


Figure 26. Mean primary task entity occlusion frequency as a function of text access technique.

A significant main effect was also found for interaction mechanization on the number of secondary task entity occlusion events across the period of a trial and found a significant main effect for text access technique [*F* (1, 69) = 38.57 *MSE* = 0.559 $p < .0001$ $\eta_p^2 = .63$]. Post hoc comparisons using the Student's t-test indicated that static interaction mechanization (*M* = 0.479, *SD* = 0.564) resulted in significantly more cases of entity occlusion than the dynamic interaction mechanization (*M* = 0.346, *SD* = 0.482).

There was no statistical evidence of that an interaction occurred between the text access technique and interaction mechanization variables [*F* (3, 69) = 0.1 *p* = .413].

Subjective Feedback

Workload

The Bedford Workload Scale (Roscoe & Ellis, 1990) decision tree procedure was completed by each participant at the end of each 600 s data collection trial. A three-way within-subjects ANOVA was conducted to compare the effect of text access technique, interaction mechanization, and initial direction for the mean reported Bedford Workload Scale rating. There was a significant main effect of text access technique on the Bedford Workload Scale rating [F (2.3, 53.5) = 5.168 MSE = 0.39 p = .006 η_p^2 = .183]. Post hoc comparisons indicated that the mean Bedford Workload Scale rating for the pop-up dialog box text access technique (M = 5.01, SD = 1.36) was rated as producing significantly less workload than both the RSVP (M = 5.42, SD = 1.61) and cinema credits (M = 5.5, SD = 1.57) techniques. There was no statistically significant difference between the times square (M = 5.3, SD = 1.61) technique and any of the other text access techniques.

Preference

Participants ranked their subjective preference for the text access techniques from 1st through 4th. One participant failed to complete the subjective preference questionnaire. Accordingly, the following is based on the input from the remaining 23 participants. Figure 27 shows preference bin frequency transformed into percent of total responses for each of the text access technique levels. There was a strong primary preference for the pop-up dialog box versus the other access techniques. When a secondary preference was considered, a preference for the times square text access technique emerged. The trend for RSVP technique reveals an increase toward the least preferred technique.

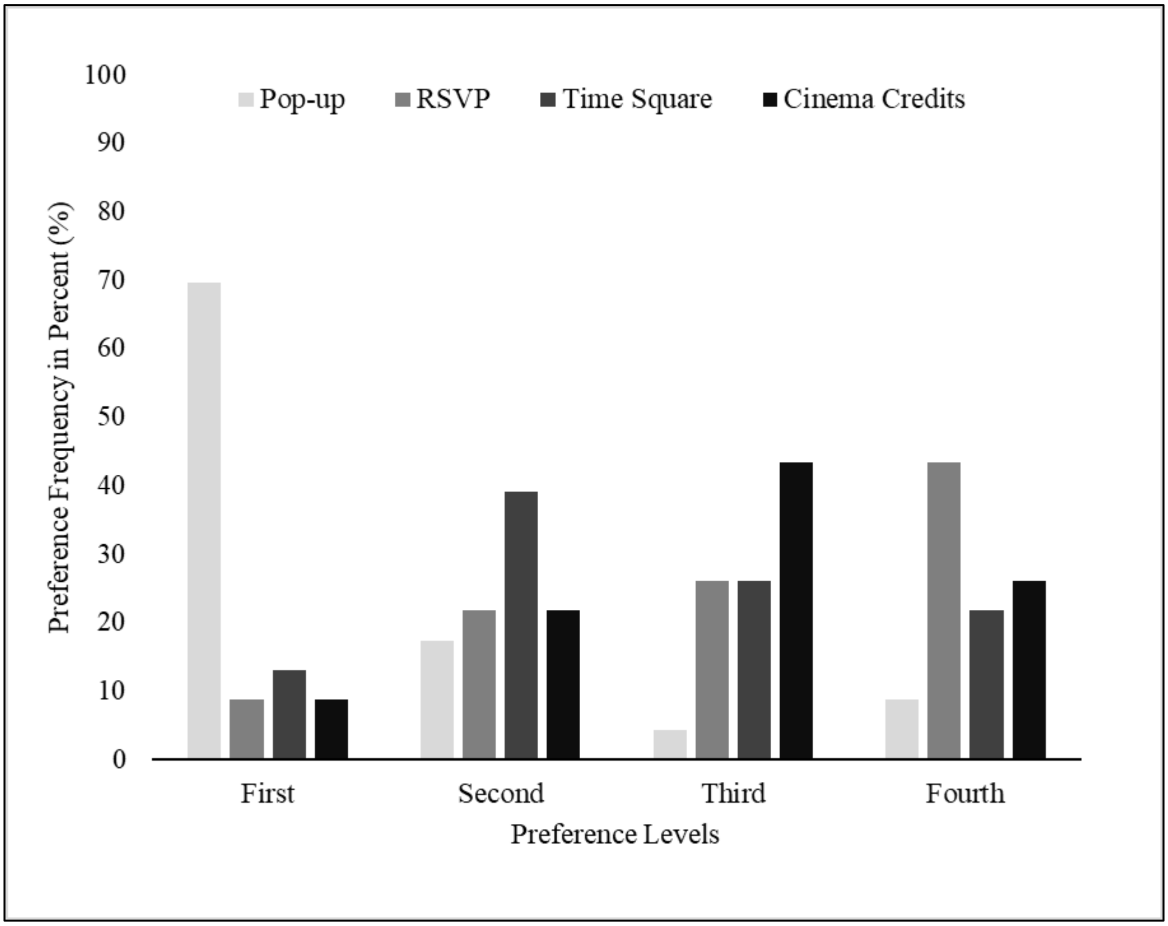


Figure 27. Text Access Technique subjective preference levels.

Mini Multifunction Display Soft Key Usefulness

Two subjective feedback questions were posed as “yes or no” responses to measure the usefulness of the spatially-constrained text access technique bezel soft keys (mini multifunction display) functionality. The first question was: “Did you find the function buttons around the entity tag to be helpful?” Of the 23 responses recorded, 16 indicated “yes” and 7 responded “no”. Thus, the positive response was 70% of the recorded input. Participant were further asked, if their response to the previous question was No, did they use the bezel soft key buttons? Six participants indicated that they did not use the buttons.

Another question was: “Specifically regarding the function buttons around the entity tag, were the "fast forward" or "rewind" buttons helpful?” Of the 22 definitive responses recorded, 4 indicated “yes” and 18 responded “no”. Thus, the positive response was 18% of the recorded input.

To investigate the use of the soft keys further, the collected data were reduced to output participant mini multifunction display soft key use during performance of the primary task. Specifically, soft key selections during primary task queries were detected, compared for accuracy (i.e., which key was selected versus the query indication), and counted. Contrary to the participant subjective feedback responses on the questionnaire, the objective data show that the soft keys were used often and mostly accurately for all of the spatially-constrained text access variants. Figure 28 shows the relative proportions of initial soft key input (i.e., the first left click on the mouse) during a query probe for all correct primary task responses. This analysis indicated that all participants used the mini multifunction display soft key functions. In fact, the majority of participants selected the correct soft key for the 9-line item as specified by the query. Participant # 22 chose to begin each entity interaction with a “start” selection. Participant # 23 used a strategy that relied almost exclusively on manipulation of the “fast forward” and “reversion” keys to initiate interaction with the mini multifunction display.

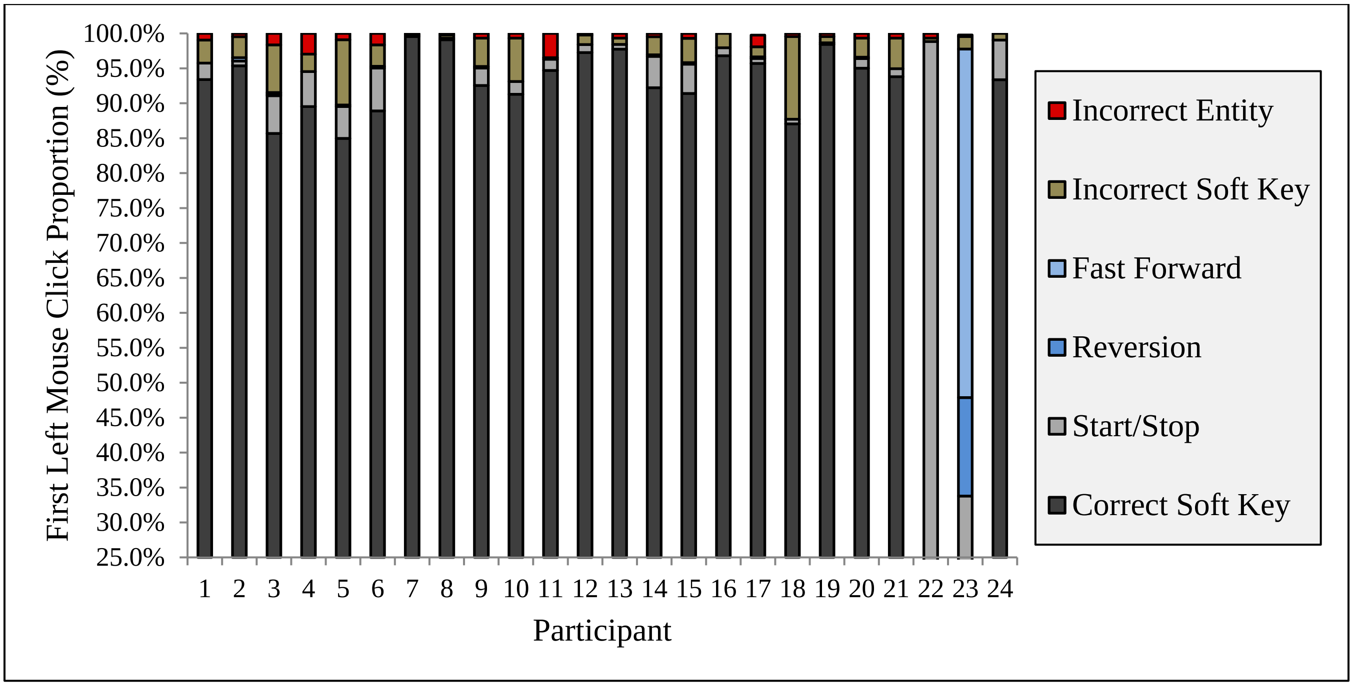


Figure 28. Initial mini multifunction left click activation or each participant.

It is difficult to explain the discrepancy between participant-reported soft key use compared to detected soft key use. Perhaps participants did not understand the questionnaire well enough to answer consistent with their actions or they may have misinterpreted the phrasing of the questions. In retrospect, an illustration to indicate what “function buttons around the entity tag” was as a reference to would have acted as a reminder of soft key use and may have helped avoid confusion. It is unlikely that participants answered in a way just to “get the questionnaire over with” because the least time consuming answer regarding soft key usefulness would have been to indicate that the buttons were helpful. Regardless, the objective performance data gave an accurate indication that 100% of the participants used the keys on every query when the text access technique included a spatially-constrained variant. Figure 29 shows the average number correct mini multifunction display soft key selections for each primary task query across all spatially-constrained text access techniques. All participants used soft keys and most used the specific key which corresponded to the 9-line item identified by the query. Again, participant #23 chose to use the “fast forward” and “reversion” soft keys to manipulate the mini multifunction display text access.

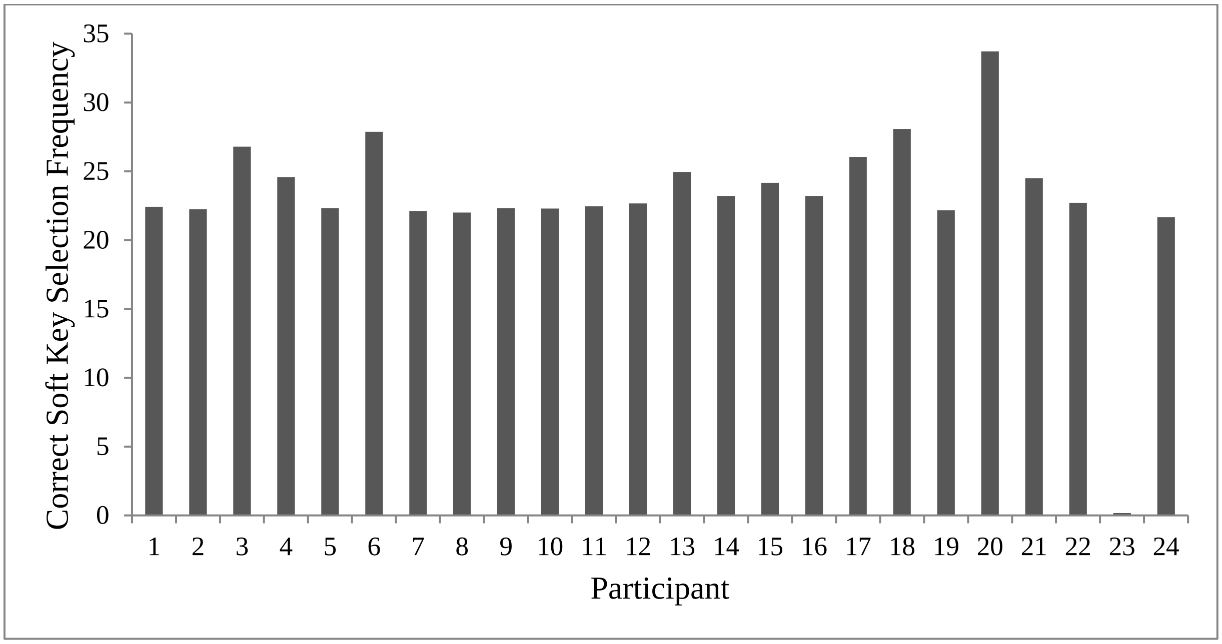


Figure 29. Correct soft key selection frequency for SC primary task queries.

Mini Multifunction Display Soft Key Response Time Performance

To further investigate the usefulness of the mini multifunction soft keys, a descriptive data analysis was performed to quantify the effect soft key use had on primary task performance compared to performance of the same task without the functionality. There were no data collected to specifically record performance of the primary task without use of the mini multifunction soft keys but reasonable values for comparison were derived from the existing data to fairly accurately characterize some comparisons. As shown in Figure 28 and Figure 29, when text presentation techniques other than the pop-up dialog box format were used for primary task input, soft keys were almost always used initially during completion of the task. When the data set was reduced to isolate the duration between the first user interaction with the entity and a correct query input, the recorded mean response time was 3.00 s (n = 10348 cases). The mean response time when incorrect input was included was also 3.00 s (n = 12440 cases). When pop-up dialog box response data were included for both correct and incorrect input, mean response time was 2.96 s (n = 16608). Finally, correct and incorrect pop-up dialog box input by itself resulted in a mean response time of 2.84 s (n = 4168). Mean response time was also 2.84 s (n = 3475) when only correct pop-up dialog box input data were included.

If the mini multifunction display soft key interface was not available, users would have to wait until the spatially-constrained text cycled to the appropriate line and value before making the query and entity comparison. The spatially-constrained text presentation was designed to complete a full cycle every 20 s and each line required approximately 2 s to complete. Given a large number of cases as were collected for the experiment, 10 s mean response time is a reasonable approximation for comparison to the Experiment 2 measured performance. The 10 s mean response time estimate is based on an assumption that the query probe task comparison was made during the first cycle occurrence and that the response to the query was essentially instantaneous.

Given the above, user primary task performance when aided by the mini multifunction soft key interface, on average, took less than a third of the predicted average time required for performance of the same task under ideal conditions without the aid of the mini multifunction soft key interface.

Summary

The purpose of these analyses was to support a valid and reliable determination of how the conceptual text access techniques and interaction mechanizations compared to one another in terms of objective human performance, subjective workload, and subjective preference. Care was taken to execute a methodology that produced data sets which lent themselves to inferential statistical analyses. The analyses themselves were performed through a systematic process including exploration, assumption testing, hypothesis testing, and results production. The overall goal was to produce generalizable findings as a contribution to the human-systems integration body of knowledge. The findings presented above are interpreted against the previously derived hypotheses in the following chapter. The objective evidence presented here as well as the accompanying interpretation formed the basis of the overall conclusions of this work in terms of contribution to the body of knowledge.

V. Discussion

Chapter Overview

This chapter is organized by first addressing the research hypotheses in light of the completed study and associated results. The specific investigative questions are then addressed followed by the identified limitations of the research. The contributions of the research are presented, followed by general recommendations.

The objective of this effort was to develop and evaluate techniques to afford operator access to relatively large amounts of text within a small graphical drawing area (i.e., a spatially-constrained drawing area). Several techniques were included based on fundamental human reading performance and prior research which aimed to provide ways to display text on small physical screens (e.g., early generation cellular telephones). The application domain for this work was text access enhancement within large area displays where it is desired that the user be able to maintain both “big picture” awareness of dynamic entities and accurate and timely access to detailed information about those entities. These are the basic performance elements of command and control tasks such as air traffic control radar monitoring or even within the cockpit itself during tactical situation display use. The conventional means for displaying text in these applications is to contain the text-based information in some sort of dialog box that is either co-located with the big picture display as shown in Figure 3, or to remotely locate the text output in a dedicated window. In either case, the operator is given access to all of the text-based information at once. There are potential negative performance implications for both of these approaches. When co-located, the presentation of the text may occlude the entities of interest within the large area display causing the operator to miss potentially important events at the big picture level. Alternatively, remote display of the text reading task may cause a similar problem because the operator is required to look away from the entities of the large area display to read the text.

This effort attempted to address the performance trade-off described above by developing and evaluating techniques for co-locating text-based information within the large area display in a way that reduced occlusion potential yet maintained accurate and timely access to the text-based information. Three different text presentation approaches were combined with a miniature multifunction display mechanism to be utilized by the operator for text-based information access. The text was provided as a continuous “feed” within a window that displayed only 12-14 characters at a time as illustrated in Figure 14 but a mechanism was developed to allow the operator to selectively jump directly to the information of interest through a mini multifunction display soft key graphical interface.

A dual-task methodology was developed to evaluate the spatially-constrained text access techniques against a conventional pop-up dialog box presentation of standardized information content. Both the primary and secondary tasks were designed to mimic the information seeking activity of a generic command and control type task where continuous monitoring of the overall entity activity was important but detailed information on-demand was prioritized and required periodically. The difficulty of the monitoring task was derived experimentally and is described here as well. An experiment was performed to objectively evaluate human performance during the primary text access task and the secondary multiple entity monitoring task. The spatially-constrained text techniques were compared to each other and to a conventional pop-up dialog box text presentation baseline technique. The remainder of this chapter will examine the study and the resulting output against the defined general hypotheses, their operational interpretations, and the concluding argument for the contributions of this work toward the scientific body of knowledge.

Research Hypotheses and Interpretation

The first general hypothesis was defined as: “Various information portrayal techniques can be devised which will result in measurable human performance differences during otherwise identical text access tasks.” Ample evidence was produced to support this hypothesis but not completely in line with the intent of how this research was structured. The focus of this hypothesis was the performance differences among the text access techniques and the baseline. The performance measures were accuracy (i.e., percent correct responses) and response time (i.e., task completion duration for accurate responses). A secondary task paradigm methodology was utilized to address this hypothesis and this associated hypothesis: “given the application of a secondary task paradigm methodology and the associated prioritization of the primary task, there will be no statistically reliable error or response time differences among the manipulation of text access techniques within the primary task.” In fact, it was found that there were performance differences among the levels of the text access technique independent variable for the primary task (query and response) as shown in Table 9.

Participants were instructed to prioritize the primary task and, within both the primary and secondary tasks, to prioritize task accuracy over response time. Care was exercised to tune the difficulty of the secondary task so that near perfect performance was attainable when the task was performed in isolation but little spare capacity should have been available for simultaneous primary task performance. The evidence indicated that participants had a difficult time either applying this prioritization schema at the onset of task performance or perhaps had difficulty applying the prioritization consistently. Within performance of the primary task, accuracy performance was, on average, better for the cinema credits text access technique than for the RSVP or times square techniques. There was no accuracy performance difference between the pop-up dialog box technique and any of the other spatially-constrained techniques. As shown in Table 9, there was less than a two percent accuracy difference among the four techniques.

There was also a primary task performance difference shown for the response time measure (task completion duration for correct responses). Here, task performance was significantly quicker for the pop-up dialog box than for any of the spatially-constrained techniques. As shown in Table 10, the times square technique was quicker than the RSVP technique while cinema credits response time was no different than either of the other two spatially-constrained techniques. The evidence also indicated that text access technique and interaction mechanization interacted with one another (illustrated in Figure 25). The interpretation of the nature of this interaction was, for the response time measure within the primary task, the pop-up dialog box technique differed more (resulted in quicker correct responses) consistently from the spatially-constrained techniques for the dynamic interaction mechanism than for the static interaction mechanization.

Taken together, the evidence described above is indicative that differences were detected among the text access techniques but the methodology was not sufficient to isolate those differences either for response accuracy within the primary task or to isolate the effects to the secondary task itself. Although, there was some evidence of secondary task sensitivity for the comparison of the text access techniques during performance of the primary task. For accuracy, the pop-up dialog box text access technique produced higher percent correct PID responses than any of the other access techniques. There was no evidence of PID task differences among the remaining three techniques (shown in Table 11).

Also related to task and effort prioritization, the following hypothesis was predicted: “given the prioritization of task accuracy versus response time, response time will not differ statistically significantly across any comparison for correct responses.” This was not supported by the evidence for either primary or secondary task performance. For the response time measure within the secondary task, the pop-up dialog box technique resulted in quicker correct responses than the cinema credits technique while the response times for the RSVP and times square techniques did not differ from either of the other two techniques. The difference between the quickest and the slowest average response time was only 90 ms while the overall average PID task response time for correct responses was 2.84 s. For the primary task, the difference between the quickest and the slowest response time was 900 ms with an overall average of 6.36 s. For the secondary task, the argument that a 90 ms difference in average response time is practically meaningful or impactful is difficult to support.

The secondary task measures were integrated by deriving a combined inverse efficiency score. The comparison between the IES and other secondary dependent measures can be seen in Table 11. The relationship among the spatially-constrained techniques and the performance measures was mostly consistent at least to the extent that the pop-up dialog box produced the best performance and RSVP and cinema credits mostly resulted in worse performance in comparison.

Subjective workload feedback was collected via the Bedford Scale (Roscoe & Ellis, 1990) as a global measure after each of the 32 trials. It was hypothesized that: “as a result the user interaction requirements of the novel text access techniques, reported subjective workload will be statistically significantly higher compared to the baseline technique but will remain in an acceptable range considering practical significance.” The overall average Bedford Rating score across all conditions indicated that, at least subjectively, participants considered simultaneous performance of the primary and secondary tasks to be challenging but tolerable (M = 5.3). Taken directly from the scale itself (Figure 24), the interpretation of this rating is somewhere between: “Reduced spare capacity. Additional tasks cannot be given the desired amount of attention.” and “Little spare capacity. Level of effort allows little attention for additional tasks.” Among the text access techniques, the pop-up dialog box condition resulted in less reported subjective workload than either the RSVP or cinema credits techniques. There was no evidence to indicate that the times square technique reported subjective workload differed from any of the other text access techniques. Thus this hypothesis was mostly supported.

Participant preference was recorded and indicated strong support for the pop-up dialog box technique in terms of its selection as the first choice when ranking the text access techniques in order of preference (Figure 27). Consistent with previous research, performance with the RSVP text access technique was not typically worse than other techniques but participants generally found it to be the least preferred among the options (Kang and Muter, 1989).

An additional hypothesis stated that: “The baseline technique which presents all information simultaneously will result in accurate and quick access to the text information but will present an occlusion problem." This hypotheses was supported. The pop-up dialog box consistently produced better performance for accuracy and response time measures within both the primary and secondary tasks. When considering the occlusion problem, the residual background and physical occlusion evidence supports the potential for occlusion to be problematic for tasks similar to those utilized here and likely for the operational tasks this study was designed to mimic. It was objectively the case that when the pop-up dialog box was displayed, it took up much more display area than that required of the spatially-constrained variants. When measured, the pop-up dialog box area difference is roughly 10 times that of the others. Even though, the effect of that difference was not clearly evident by an associated reduction in performance of the PID task among the collected performance experiment data. This could have been because there were too few entities active in the scenario to result in enough an occlusion problem to negatively impact the PID task. Also, the quickness with which the primary task could be performed also reduced the amount of time any occlusion was present and able to cause participants to miss a PID event.

The secondary task occlusion area analysis was performed to measure the frequency with which a PID entity was occluded by 75% to 99% when an "I" or "D" change event occurred. The analysis clearly provided evidence that occlusion occurrence was significantly more likely for the pop-up dialog box technique versus the spatially-constrained techniques (Figure 26) even though an average of 4 entities were present on the screen in this experiment. Nevertheless, the following specified hypothesis was not supported by the evidence: "The increased display area required of the baseline text access technique will result in a statistically reliable decrease in secondary task accuracy performance due to the occlusion of events of interest compared to the novel designs." Again, the evidence does support the concept that the potential for missing events due to occlusion existed and may therefore become a significant problem in displays with a higher density of entities or even more complex text content that takes more time to access or utilize.

The hypothesis which stated: “due to the prioritization of accuracy for both the primary and secondary tasks, response time for correct responses will not differ statistically significantly across text access techniques” was not supported by the evidence. As discussed earlier, the ability of these participants to consistently maintain applied task and performance priority was not evidenced by the findings. For that reason, the possibility that these data are affected by accuracy being traded for speed is a distinct one but since this study was conducted as a within-subject, repeated measures experimental design, it is unlikely that prioritization strategies shifted in any systematic manner within individuals across their experience with the tasks. The methodology employed here is likely representative of any real world population performing similarly challenging tasks without the benefit of extensive training and consequence based motivation.

The following hypothesis is likely true even though the evidence did not support it directly: “A portrayal technique which occludes information will interfere more with tasks that depend on the use of the occluded information compared to novel access techniques that have been designed to minimize occlusion.” It is possible that the lack of evidence in support of this hypothesis was a result of the number of entities presented in the experimental scenario. There were simply not enough opportunities for occlusion to show a reliable performance decrement. The following hypothesis speaks more to the potential for occlusion to be a performance driver: “The measurement of occlusion during performance of the experimental tasks will result in statically significant higher potential for secondary task error during use of the baseline text access technique compared to use of the novel designs.” The area analysis performed showed clear evidence that the potential for secondary task effects due to occlusion were significantly more likely for the pop-up dialog box text access technique compared to any of the other access techniques which were designed to avoid that potential.

This next hypothesis addresses the interaction mechanization independent variable manipulation: “a novel design feature which affords text access to remain physically closer to dynamic events of interest will generate advantageous performance compared to that which allows the separation of the events of interest from the text access task.” There was little evidence to support that the interaction mechanization had much of any effect on task performance or subjective feedback. There were no main effects of interaction mechanization among the accuracy or response time data within the analyses. On the other hand, as previously mentioned, there was an interaction for the primary task response time measure between text access technique and interaction mechanization. The nature of the interaction provided evidence that dynamic interaction mechanization resulted in the quickest response time for the pop-up dialog box condition compared to the other three text access techniques. Alternatively, for the static interaction mechanization condition, the pop-up dialog box text access technique did not differ from either the times square or cinema credits text access techniques (Figure 25). Additionally, there was evidence to support this hypothesis in the secondary task entity occlusion findings. The dynamic interaction mechanization condition resulted in significantly fewer occlusion events than the static interaction mechanization condition. According to the earlier interpretation of this measure, there was more potential for secondary task PID error for the dynamic versus the static interaction mechanization.

Since there was no direct performance benefit support attributed to interaction mechanization within the secondary task, the following hypothesis was not evidenced: “Text access attached to the dynamic entities will result in statistically significantly less secondary task error performance when compared to text access fixed in screen coordinates.”

The hypothesis which indicated: “the manipulation of text access separation will not result in a statistically significant difference in reported subjective workload” was supported. There was no evidence of a subjective workload difference attributable to that manipulation of interaction mechanization.

Investigative questions

Question 1: What is the best way to present the most amount of usable text-based information within a constrained amount of display real estate?

Ideally, the evaluation of the selected text access techniques would have indicated a clear performance advantage of one of the novel presentation types over the remaining two. Further, the nature of that evidence would have shown a clear performance advantage within the secondary task in the absence of either response time or accuracy difference elsewhere. In reality, based on the results of the overall comparison, addressing this question must be more nuanced. For secondary task performance, the evidence indicated that there was no statistically reliable difference among the spatially-constrained text access techniques but, accuracy was best for the pop-up dialog box condition. This is counter intuitive if the accuracy difference was attributable simply to occlusion of the secondary task entities during PID events. Alternatively, secondary task accuracy was negatively affected by both the time and attention required of participants to perform the primary task with the spatially-constrained text presentation designs and their associated access interface. While it was the case that occlusion occurred more often for the pop-up dialog box condition versus the others (Figure 26), it is likely that the occlusion occurrence duration was not long enough to negatively impact secondary task PID accuracy performance.

In terms of subjective workload, there was no difference between the times square condition and any of the other techniques while both the RSVP and cinema credits conditions were shown to generate significantly higher rated workload than the pop-up dialog box condition. Another potential discriminator was the primary task accuracy measure itself. Statistically, the cinema credits access technique produced better performance than the remaining spatially-constrained text presentation techniques. Pop-up dialog box primary task accuracy did not differ from any of the other techniques. Beyond the statistical result of this comparison, the mean difference between the best performance and the worst performance was less than two percent. Again, no definitive “best” performer was revealed.

Considering response time for the secondary task, the slowest mean performance was recorded for the cinema credits condition compared to the pop-up dialog box text presentation technique. The remaining two techniques did not differ from any of the other conditions. Within the primary task, among the spatially-constrained techniques, the quickest mean response time performance was recorded for the times square condition while the slowest response time was produced by the RSVP condition.

Among the spatially-constrained techniques, there appeared to be a preference for the times square text presentation technique over the cinema credits and RSVP techniques (Figure 27).

There was little evidence to support a clear differentiation between the static or dynamic levels of the interaction mechanization manipulation. Although, the objective display area analysis showed significantly more potential for entity occlusion for the static condition versus the dynamic condition. For the primary task response time measure, interface mechanization interacted with text access technique in a way that showed better performance for the pop-up dialog box technique for the dynamic condition versus the static condition. This lends some support for a dynamic interaction mechanization benefit.

Given the evidence, it can be stated with confidence that participants were able to perform the primary task comparatively well with all of the text access techniques. Among the spatially-constrained techniques, when considering primary task accuracy, the cinema was the “best” text presentation technique. For overall performance (all tasks and measures), the times square and cinema credits techniques appear to narrowly but consistently edge out the RSVP technique which consistently seemed to produce among the poorest performance.

The objective display area analysis provided clear evidence that, depending on the complexity and number of “big picture” entities requiring monitoring within the operational scenario, the spatially-constrained text access approach is useable and will help to reduce occlusion. If occlusion increases the likelihood of performance decrement, then the spatially-constrained text access approach is less likely to result in a performance decrement due to that occlusion.

Question 2: How best should operators be given access to the content of text-based information when the spatial dimension for its presentation is constrained?

This question was intended to address the interaction required of participants to satisfy the query probe questions. Participants were asked to comment on their use of the mini multifunction display soft key functionality as well as any relative performance differences between the interaction mechanizations (static versus dynamic).

Participants were asked: “Did you find the function buttons around the entity tag to be helpful?” Further, participants were asked: “If yes, how and why?” To the first question, 70% of the respondents indicated that they found the mini multifunction bezel functions useful. When asked how and why they found the functions useful, the feedback was in line with the intended functionality of the interface.

Participants used the soft keys to advance directly to location of the standard 9-line information as defined by the query probe. Appendix C includes the full recoded output of the questionnaire (Appendix B). The objective analysis of button use revealed that the participants actually used the soft keys often and accurately (Figure 28 and Figure 29).

Where it was indicated that the soft key functionality was not useful, the comments mostly addressed a lack of utility among the fast forward and reversion functions. Since these functions were specifically included in the questionnaire, more insight into their usefulness was available. Specifically, the following question was asked: “Y/N? Specifically regarding the function buttons around the entity tag, were the "fast forward" or "rewind" buttons helpful?” The comments indicated while the functions were occasionally used, and that they performed the desired function, the buttons which enabled direct relocation to a specific line item were more useful and served much the same function. There were no comments to indicate that the mini multifunction display soft key functionality was better for any of the spatially-constrained text access techniques versus the others. Of course there was no reason to use the soft key functions for the pop-up dialog box baseline condition.

As discussed earlier, there were few objective performance measures which differentiated the interface mechanization conditions to show any clear advantage of one versus the other (static versus dynamic) for the experimental task. Participants were asked specifically about the comparison via the following question: “Y/N? In some cases, while answering the queries, the text tag remained fixed in place on the screen while other times it continued to move with the entity. Do you think this affected your task performance in any way?” The response indicated that 74 percent of the participants felt that performance was affected in some way (Appendix C). The following comment was selected as a good summation of the overall feedback: “If the text stayed in place, I found it would get in the way of watching the entities. It was also another task to have to click on it to get it to move again which drew my attention away from the main tasks." In general, participants consistently indicated that, when static, performance of the primary task benefited but that secondary task monitoring was less disjointed during the dynamic condition.

Overall, the findings were consistent: interaction with the spatially-constrained text access techniques was acceptable and no major modifications are required to otherwise significantly enhance performance. The soft keys which provided direct access to specific lines of text-based information were credited with being the most "helpful" from a usability perspective compared to the other features. In relative terms, the fast forward and reversion functionality was reported to be marginally useful. These soft keys could be used for other purposes if desired without much capability loss. Similarly, the interaction mechanization choice should perhaps be based on whether the system designer places emphasis on a big picture monitoring task (the secondary task for this study) or a text access task (the primary task for this study).

Question 3: What are the relative advantages and disadvantages of the various mechanizations within an operationally representative task environment?

The objective of this research was to develop and evaluate techniques which afford a user access to large amounts of text within a small display drawing area (spatially-constrained). Beyond the text presentation techniques themselves (times square, cinema credits, and RSVP), access functionality similar to mini multifunction display soft keys was developed to give users direct access to specific points within the body of the text. Progress metering was included as well. Where direct comparisons were made between the use of the pop-up dialog box text presentation versus the spatially-constrained techniques (such as during the primary task), the pop-up dialog box did not definitively out-perform some of the others. For the percent correct responses measure, the statistical evidence indicated that performance with the pop-up dialog box access technique did not differ from performance using the spatially-constrained techniques. It was not surprising to find that response time was quickest with the pop-up dialog box condition but the mean difference between that fastest response time and the slowest (RVSP) was still less than one second. Given the typical real-world priority of accuracy over speed, the study evidence supports a conclusion that practical performance among the text access techniques was quite similar. Given the fact that the physical design of the spatially-constrained text access techniques resulted in significantly less potential occlusion of the overall display area, those techniques seem a viable option to be considered by interface designers when the occlusion threat of a pop-up dialog box needs to be avoided or should be mitigated.

Among the spatially-constrained techniques themselves, the practical differences for percent correct performance and response time were also small. In general, where there were differences; the RSVP design tended to produce the worst objective performance and the worst subjective preference. For primary task percent correct responses, the cinema credits technique was best and for the response time measure, there was little difference between the cinema credits and times square conditions. Unless there is a compelling reason to consider it otherwise, RSVP as a text presentation technique should be avoided.

Question 4: How well does the methodology measure the human performance differences resulting from the variations among the text access designs and information portrayal features?

The subjective workload feedback gathered from the participants indicated that the experimental tasks were challenging but, overall, performance was good. The intent of the secondary task development effort was to generate a sensitive measurement instrument by producing a task that was effortful but permitted participants to produce almost perfect performance when completed in isolation. The evidence supports that this was likely achieved albeit, participant task prioritization may not have been as consistent as desired. Although some differences were detected among comparisons of the various text access techniques, performance attributable to occlusion of the secondary task entities was not straight forward. For this experiment, a larger number of entities may have remedied the issue by generating a scenario where the potential for occlusion would have likely increased. But, if the number of entities was increased, other secondary task element adjustments such as decreasing the number of PID events would have been necessary to maintain the level of difficulty at the current desirable level. Any increase in difficulty would likely have resulted in an unacceptably difficult secondary task. Because response to the PID events was the performance measurement basis of the secondary task, an emphasis on including a reasonable number of the events at a reasonable event rate was placed over the number of entities in the scenario. Given these considerations, if the number entities were increased and the PID event rate was reduced, or maybe even tuned to match a real-world system event rate, trial duration or the number of trials would need to be increased to produce the same quantity of measurement data as the current experiment. It is not unreasonable to do this and this is especially true where fewer user interface formats are included for comparison (i.e., a baseline versus one novel format). Experiment 1 demonstrated that this trade-off is reasonable (Figure 11).

Performance with the techniques turned out to be similar and the display area objective measures proved that the pop-up dialog box technique did occupy more display area across completion of the tasks compared to the spatially-constrained access techniques. For this methodology, it was critical that participants exercised the proper task completion prioritization to maintain maximum task performance measurement sensitivity. The fact that primary task performance differences were recorded among text access techniques is evidence that task prioritization discipline was not well maintained. Under conditions of real-world task prioritization, it would be unlikely that this would be the case. For future use of this secondary task methodology, perhaps more care should be utilized to ensure that proper task prioritization is maintained.

Limitations of the Research

The experimental setup and associated tasks were designed to be representative of operational tasks where the evaluated text access interfaces were intended for application. This was true in a generic sense but there was no particular system on which the experimental methodology was based. Instead, the intent was to be able to more widely generalize the findings across a variety of systems and users. The experimental design was also developed so that no special skill or subject matter expertise was required of the participants utilized for data collection. In terms of limitation, the decision to pursue the level of operational representation used for this study resulted in the ability to make some general design recommendations based on the text access technique comparisons but further evidence is required to confidently apply specific interface mechanizations within an actual operational system. This is not unusual at this level of technology maturation.

One example of a generalization limitation is the basis of the text content used for the study. As previously discussed, the spatially-constrained access techniques designed for this evaluation likely requires that the body of text be presented in a format with a standardized sequence. The findings here should not be extended beyond the standardized format restriction in the absence of an evaluation designed specifically for some other application. Similarly, the number of soft keys used for the study was well suited for text based content with ten elements. If the number of text elements differ in number or symmetry, the number of keys, their size, and their operation should likely change but the impact of the change is difficult to predict.

The use of an operationally representative task and naïve participants resulted in a challenge toward achieving the desired prioritization of the primary task over the secondary task and for accuracy over response time. The evidence collected for this study indicated that this was likely problematic. The interpretation of the outcome was certainly less straight forward than it would have been otherwise.

Another area of limitation was the relatively low number of entities utilized for the secondary monitoring and reporting task. A goal of this research was to investigate the potential to minimize occlusion cost of display area occupied by the pop-up dialog box compared to the spatially-constrained text access interface design. The “big picture” monitoring task which had the potential to be negatively affected by occlusion was tuned to maintain an average of four entities. In retrospect, and compared to real-world operational applications, both the number and size of the entities which were monitored for the secondary task were somewhat resistant to any occlusion based decrement. Figures 1, 2 and 3 illustrate the relatively large number and small size of big picture entities monitored in real-world applications while Figure 19 is a representation of the entities for the monitoring task used for the experiment. In Experiment 2, had significantly more and smaller entities been used, it is possible that the task may have biased the results toward an advantage for the spatially-constrained text access designs. A remedy may be testing the specific designs within real-world application representative scenarios with real-world system users. This is a conventional process of technology maturation and should likely include only a small selection of format comparisons.

The manipulation of the static versus dynamic interaction mechanization was intended to exercise a comparison of detailed text presentation related to an entity of interest (one that required monitoring) at a location remote from that entity (static) versus one co-located with that entity (dynamic). Examples of both of these approaches can be found among conventional designs within operational systems (Figures 2 and 3). There is a third approach which was not tested where detailed or drill-down text based information related to a selected entity is displayed in an area that is outside of the scenario “gaming area” but close by as a window specifically dedicated to the display of text based information. This “remote” presentation technique is designed to afford text access with no potential occlusion impact to the big picture monitoring task. Still, there is a potential monitoring task performance cost associated with the need for the user to shift foveal line-of-sight to the remote text area away from the entity monitoring task. Figure 30 is an illustration of text-based detailed entity information presented around the periphery of the big picture monitoring area. Since the objective of this study was focused on the potential occlusion effects where drill-down and big picture attention is of simultaneous interest, the remote text access interface variant was not investigated. The static interface mechanization condition was most similar to the remote text access concept but did not require the operator to fixate outside the display area for the secondary task.

The limitations identified for this research are not atypical or unique to this type of early TRL study. It was necessary to make some trades and compromises so that a manageable and appropriate scope was maintained for the data collection effort. In general, the limitations discussed here did not significantly reduce the value of the research contribution or negatively influence the level of confidence with which design recommendations are offered based on the study findings.

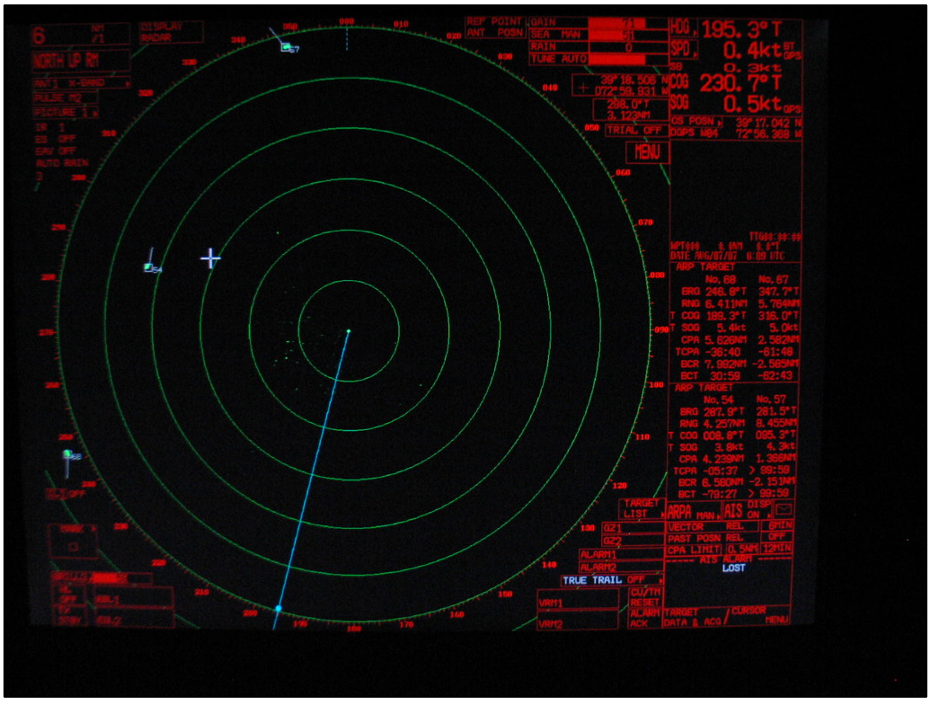


Figure 30. Example of remote drill-down text access technique (Frantzdale, 2007).

Research Contributions

Summary

The following paragraphs outline the contributions the completion of this research offers to the scientific and engineering body of knowledge. The introduction chapter of this dissertation outlined the contributions of the previous research used as the building blocks upon which the designs and methodology for this study were derived. Ideally, the knowledge and suggestions generated by the completion of this study are useful in further defining the design elements and application of the concepts and techniques evaluated here. The contributions include the text access interface concepts themselves, recommendations for concept refinement based on the limitations of the designs investigated here, suggestions for operational applications, and the validation of the methodology developed for the evaluation.

Contribution 1: Text access design concepts evaluated objectively within an operationally representative scenario environment:

The experimental methodology developed for this research was designed to be representative of the operational application where a user is required to simultaneously be aware of the state of multiple dynamic entities while periodically being required to “drill down” and access detailed information about individual entities for purposes of decision making. In general, the tasks were developed to be representative of the fundamental types of information processing elements which are common among most command and control information sources. The unique design of the spatially-constrained text access and presentation techniques was intended to promote enhanced user performance for the defined tasks, i.e., detail information on demand and continuous global monitoring in terms of accuracy for a given task duration. The evidence produced by this study supports the notion that the spatially-constrained interface concept afforded performance mostly on par with the more conventional pop-up dialog box text access technique while utilizing less display real estate to do so. The evidence was consistent both objectively through application of performance data and display area measures, as well as subjectively based on workload rating and participant feedback. Subjective preference for the pop-up dialog box was fairly consistent but it is also the case that participants have had much more actual experience with this style of interface approach versus the novel spatially-constrained text access concepts. From this perspective, the spatially-constrained text access concepts can be offered to the interface designer as a viable option when there is an operational requirement to minimize the display real estate required to present multiple word text in an acceptably readable way while avoiding the potential for problems caused by occlusion of other graphical display elements.

Within the spatially-constrained text access techniques, the objective and subjective preference evidence indicated that the RSVP presentation method should be avoided unless separate performance data are generated to support its use. For this study, there was no compelling reason to indicate that RSVP be utilized given that the space allotted for text presentation was held constant across all three of the evaluated methods.

Contribution 2: Recommendations and limitations for concept refinements for operational applications (including format technique mechanization, features, and interactions):

For this research, the novelty of the text access user interface design was more the mini multifunction display functionality than the text presentation techniques, progress feedback feature, or interaction mechanization. The comparative evaluation of the combination of these manipulations was the focus of this work but the elements other than the mini multifunction display design had been the basis of previous research for various other applications. As a concept, the mini multifunction display was developed out of the necessity to support a way for users to navigate the content of the text-based information while maintaining the desired minimal display real estate objective. Primary task performance when aided by the mini multifunction soft key interface, on average, took less than a third of the predicted average time required for performance of the same task under ideal conditions without the aid of the mini multifunction soft key interface. Objective analysis of mini multifunction soft key use revealed that participants relied heavily on the functionality to perform the primary task.

As discussed earlier, a limitation of this approach was the fact that the underlying text-based information must be sequenced in a standardized fashion for successful application of this approach. A related limitation was that the user needed to either be familiar with the structure of the standard sequence or some reference was necessary to enable the utility. For this research, the 9-line structure (Table 3) was selected which led nicely to the use of five function keys on the top and five function keys on the bottom of the symmetric multifunction display structure (Figure 17). In terms of refinement, the implications of the use of more, fewer, or a non-symmetric soft key layouts are not known explicitly but there is little reason to believe that fewer function keys would be problematic. On the other hand, the addition of more function keys may pose a challenge as the area dedicated to each button would be reduced, potentially hindering the user’s ability to select the desired buttons. The implications of using more than 10 function keys (five on the top and five on the bottom) should be determined empirically. Other modifications such as display resolution, interaction mechanization, text content complexity, frequency of use, etc. may have a significant impact on what the maximum acceptable number of usable function keys may be.

Another function key feature utilized for this study was the graphical expansion of the function key set when the mouse cursor was hovered over any part of the soft key set (Figure 17). Within the expanded soft key set, an individual key was highlighted when the mouse cursor was hovered over it. The reason for this functionality was to minimize the size of the text access graphics while enabling acceptable ease and accuracy of key selection by the user. For refinement consideration, the individual function key could be expanded instead of the entire set. This would eliminate the need for the highlighting feature within the soft key set as well as lend itself better to use of an asymmetric set of function keys within the mini multifunction display.

The current study did not include any type of labeling for the soft keys but that would have been potentially helpful and certainly may be necessary in other applications. Since one of the identifiers used for the query probe was line number (Figure 19), selection accuracy and response time performance may have been enhanced had the specific line number associations been explicitly included within each corresponding soft key. Exactly how that labeling is rendered is another design question. Should the number labels be there full-time with the miniature keys prior to expansion or would that present an unacceptable clutter issue? Alternatively, should the labels only correspond with expansion? If that is the case, does labeling remain useful?

Another area for potential refinement is the approach used for direct selection, pausing, stopping, fast forwarding, and reversion. For this study, only the basic functionality was utilized for simplification purposes. The sophistication of these dynamic navigation features could be increased significantly with the potential to enhance text access performance. Depending on the information content of the text, direct selection, fast forward, and reversion functionality could be designed to advance directly to the data value instead of to the beginning of each individual 9-line string. This would have the potential to reduce the response time for tasks similar to the query probe used for this study.

There are myriad ways small refinements may be applied the mini multifunction concept developed for this research. The examples presented here are not intended to be an exhaustive list of options nor would it be appropriate to attempt to do so within the scope of this dissertation. This research stands as objective evidence that the mini multifunction display soft key concept was shown to have merit toward providing quick user access to text-based detailed information in a spatially-constrained manner. Refinements to the functional elements of the concept should be considered within the operational context of employment and the information content of the underlying text.

Where refinements are considered for future operational utility, this research offers a methodology for user performance evaluation. Utilizing the methodology developed here for future evaluations will also act to further validate it as a useful research tool. This idea will be expanded upon below within the "Contribution 4" discussion.

Contribution 3: Example application suggestions:

Applications for the concepts developed for and evaluated by this research are numerous. The application focus for the study itself was an operational environment where a user is simultaneously interested in the "big picture" relationship among multiple dynamic entities and detailed "drill down" information pertaining to specific entities. A command and control environment such as air traffic control area radar display monitoring was offered as an operational example. Beyond that type of application, the spatially-constrained text access concept potentially has a much broader application. The practical affordance of the concept is access to a large amount of detailed information in a small space. The fact that text was used for the present study is not an actual limitation of the concept. There is no reason to believe that access to graphical, symbol-coded, or imagery-based information could not be the presentation content for which the access user interface is applied. The higher level utility of the spatially-constrained concept is to afford access to large amounts of information in a relatively small space where there is a desire to minimize clutter and occlusion.

One application suggestion is use within the portal based large area display (LAD) user interface incorporated on the flight deck of the latest generation aircraft. Within a 5th generation single-seat tactical aircraft, the portal based LAD is utilized but drawing space on the display surface remains at a premium due to the large amount of information that is available for presentation. Presently, drill-down information related to specific points of interest is made available to the pilot via a dedicated area of a portal as a dialog box which often includes all of the information at once. This is true even if the pilot is only interested in one specific bit of information. It is suggested that a spatially-constrained technique such as developed here could offer the pilot access to that same information at less display area cost with similar or potentially superior efficiency. This is especially true when the visual user interface is coupled with a direct selection capability that lends accuracy and speed.

Another display space concern is the drawing surface of the helmet-mounted display (HMD) field-of-view (FOV). The HMD is a transparent display where critical information is presented in way that the pilot can maintain visual contact with points of interest in the out-the-window world. Because the pilot cannot "see around" HMD presented information that is FOV stabilized, unnecessary clutter should be strictly avoided. For this reason, text based or other grouped information utilizing dialog boxes is not typically presented on the HMD. If the content of that information is desired, the pilot must look back into the cockpit to access that information on the LAD. This behavior thus causes the pilot to at least temporarily shift visual attention and physical visual accommodation from the outside world to the LAD and then back again. This undermines the reason the HMD exists, i.e., to maximize "head up" and "head out" pilot attention. Perhaps HMD based information access can be afforded and become acceptable given a properly designed and mechanized spatially-constrained access technique.

The suggestions offered above are just two among a large number of potential applications where there may be a performance benefit represented by the use of spatially-constrained information access techniques. There are likely potential advantages anywhere there is a desire to minimize the space required to present standardized information content.

Contribution 4: Development and validation of an evaluation methodology and analysis strategy which can be adapted to a wide variety of C^3I representative human/machine interface objective evaluations.

The methodology developed for this research has the potential to be reused for evaluations where the objectives of the novel user interface are similar to that of the spatially-constrained text access techniques compared here. That is, there is a baseline user interface for comparison and the operational application includes monitoring and information selection tasks. Of course the limitations of this research as discussed previously should be taken into consideration for other evaluations but, nevertheless, the utility of the methodology was shown to have merit. Similarly, the human performance analyses utilized here proved to be straight forward and logically interpretable. Chapter III of this dissertation provides a manuscript that was written specifically to address this evaluation methodology contribution.

Contribution 5: Communication of the research to the scientific and engineering community.

To date, this research has produced two separate submissions for peer-reviewed publication. The first manuscript titled "Development and validation of a secondary task environment for assessing visual-psychomotor tasks" was submitted to *Theoretical Issues in Ergonomics Science* on 28 March, 2019 (Geiselman, Heft, & Miller, 2020). The first peer review disposition led to a revised and resubmitted manuscript (4 May, 2019). The revision was completed and submitted on 14 May, 2019. The first revision of the manuscript was accepted for publication on 6 February, 2020. This manuscript makes up the majority of the Chapter III (Experiment 1) content of this dissertation. The submission was intended as a description, evaluation, and validation of a methodology that can be used as the basis of a sensitive secondary task for research similar to that presented as Experiment 2 in this dissertation.

A second manuscript was prepared in response to a January 2019 call for submissions for the IMAGE 2019 Conference held in Dayton, Ohio on 25-26 June, 2019. Submission proposal acceptance was received 1 March, 2019 with a first draft of the manuscript due for review on 30 April, 2019. Review comments were received 15 May, 2019 and the final manuscript was submitted on 30 May, 2019. The content of the paper was presented at the conference on 25 June and the paper was subsequently published in the conference proceedings (Geiselman & Miller, 2019).

The content of the IMAGE 2019 submission included a subset of the full data set collected for Experiment 2 of this dissertation. The paper included data collected for 9 participants (versus 24 participants included for Experiment 2) and the evaluation comparison was limited to the RSVP and DB text access techniques as well as a manipulation of DD and SI interaction mechanizations. The manuscript is presented in its entirety in Appendix E.

The next submission planned is a manuscript prepared to include all of Experiment 2 from this dissertation. The intended publication is the *Displays* journal.

VI. Conclusions

The evidence produced as a result of this research supports several conclusions. First, techniques which afford user access to text based information can be designed which minimize and perhaps optimize the amount of physical display real estate required for that access. Specifically, the mini multifunction display soft key interface evaluated for this research produced performance as good as or nearly as good as conventional techniques where all of the text is displayed at once within a much larger area. The larger area required for the conventional dialog box format represents the potential to cause occlusion problems compared to the spatially-constrained designs. Within the spatially-constrained access techniques, scrolling text from right to left or from bottom to top as a portrayal method produced similar performance. Thus, the selection for which is best to use can be left to the designer or user to decide. The evidence consistently supported the recommendation that the use of Rapid Serial Visual Presentation (RSVP) should be avoided. Not only was performance mostly worst when this presentation technique was used, the application of this technique was undesirable based on the subjective preference. This is also consistent with findings within the RSVP research literature (Muter, 1996). The comparison of static versus dynamic interaction mechanization was mostly inconclusive based on the performance measures used for this study. Here also, perhaps whichever approach is most well suited for the application should be used. Lastly, there remains potential performance gains to be realized for the overall concept in the form of refinement of the access and presentation designs based on the intended target application.

The methodology developed for this research shows promise for future use but it is not without its limitations. Modification of the methodology to more closely resemble a specific application of interest is encouraged at least in terms of number of entities included in the monitoring task. Care must be taken to tune the parameters of the task such as event rate to ensure that the resulting workload is representative of the application of interest or at least to ensure that the workload is acceptable. Also, to the extent possible, participants trained to maintain consistent task performance prioritization will also help produce the best objective performance-based conclusions.

Appendix A: Participant Demographics Questionnaire

DEMOGRAPHIC QUESTIONNAIRE

HUMAN PERFORMACNE STUDY
GEISELMAN, MILLER, HEFT, & MENKE
AFRL/711 HPW/RHCV and AFIT/ENV

Participant number ____ Age _________

(circle one):	20/20 Vision	Corrected to 20/20	Less than 20/20
(circle one):	Color Vison	Normal	Deficient
(circle one):	Glasses	Contacts	Neither
(circle one):	Left-handed	Both	Right-handed

Hours per week playing video games: _____

If applicable, what type of video games do you play? ________________________________

If applicable, what type of video game system do you use? (circle all which apply):

Computer TV Portable

Hours per week watching TV: _____

Hours per week on a computer: _____

Is computer use for (circle all which apply):

Work Leisure Gaming

Comments:

Appendix B: Participant Post-Test Questionnaire

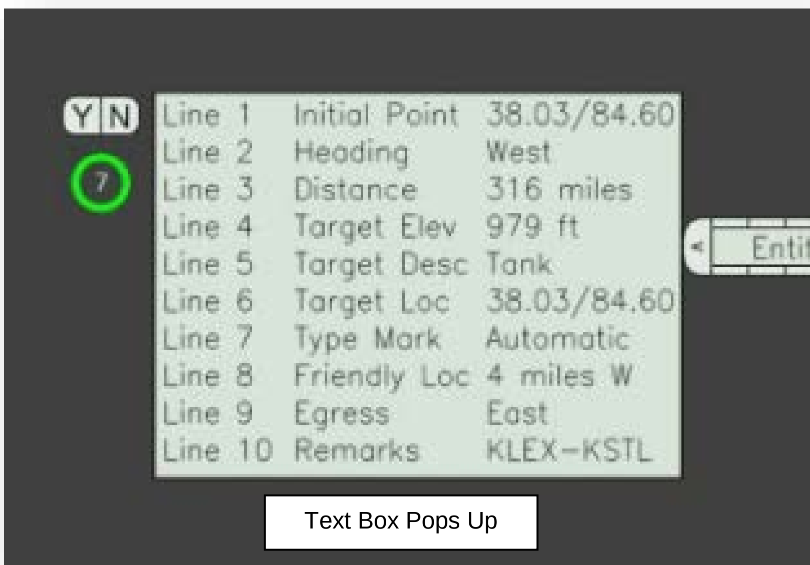
Text Access Study
Post-Participation Questionnaire

Participant ID: ____________________

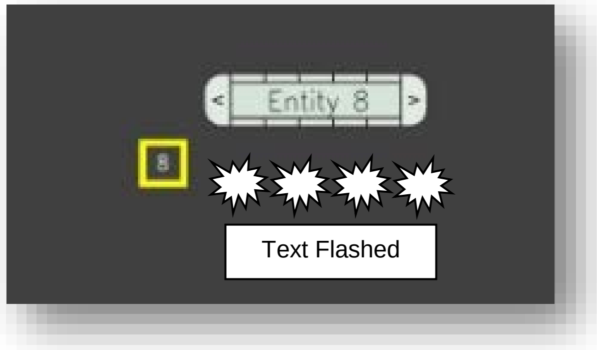
Date: _____________________

Display Variation Concept Reference

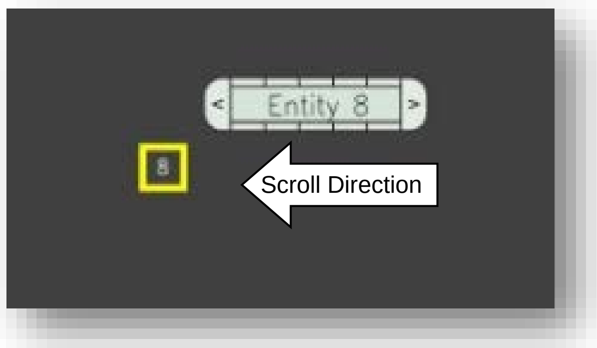
Concept ID (**A**): Text presented as a pop-up box



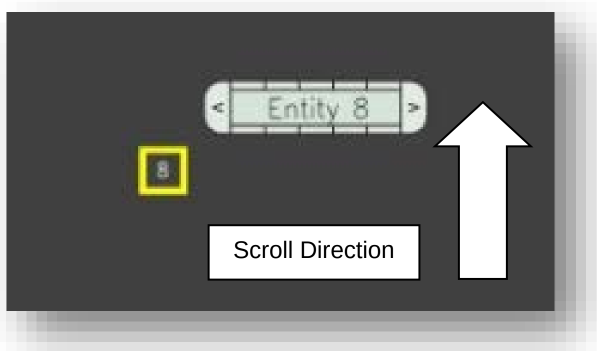
Concept ID (**B**): Text presented as a group of characters flashed in sequence.



Concept ID (**C**): Text presented as scrolling from right to left.



Concept ID (**D**): Text presented as scrolling from bottom to top.



Questionnaire Responses

1. Using the concept IDs above (**A**, **B**, **C**, and **D**), rank the text presentation types in order of your preference. For equal preference, list concepts on the same line below.

   First preference __________

Second preference ________

Third preference ________

Forth preference ________

Where a preference among the text presentation concepts is shown, please indicate what you found preferable and/or objectionable as a comment below:

2. In some cases, while answering the queries, the text tag remained fixed in place on the screen while other times it continued to move with the entity. Do you think this affected your task performance in any way (please circle **Yes** or **No**)? If you answered **Yes**, please indicate below **Why** and **How**?:

3. Did you find the function buttons around the entity tag to be helpful (please circle **Yes** or **No**)? If you answered **No**, did you use them at all (please circle **Yes** or **No**)? If you answered **Yes**, please indicate below **Why** and **How**?:

4. Specifically regarding the function buttons around the entity tag, were the “fast forward” or “rewind” buttons helpful (please circle **Yes** or **No**)? If you answered **No**, did you use

them at all (please circle **Yes** or **No**)? If you answered **Yes** to either, please indicate below **Why** and **How**?:

5. Please indicate below any general comments you have regarding your experience with this study and/or suggestions for improvements for the text presentation concepts.

Thank you very much for your participation!

Appendix C: Participant Post-Test Questionnaire Output

Where a preference among the text presentation concepts is shown, please indicate what you found preferable and/or objectionable as a comment below:	**Participant Comments:**
Key: A = Pop up Dialog Box B = RSVP C = Times square D = Cinema Credits	I liked being able to see all of the data at the same time. I would often miss pieces of information on options other than A.
	I felt C was the least intrusive to me while trying to answer the questions while still tracking other objects.
	My preference was the pop up box because it allowed my eyes to focus on the answer and not second guess what I actually saw when comparing it back to the query at the top.
	Like A because it was static-easy to read - but realized that I was blocking part of the screen and missed changes in other irons - this was ok in the end because I was supposed to prioritize that "text" task. Liked B because the text was more static (even for the brief time it was up). Next D - better on the eyes I thought over C.
	I preferred preference C as it felt more natural to read from left to right. Option A left too much space to look at and I found it more difficult to find the correct line.
	A was the easiest to me, because the information was all right there. C and D did not really give me any trouble but on some, like the coordinate lines, B was too fast for me.
	A, while the most restricting in terms of available space allows two benefits; the quick and reference-able access for comparison because it does

	not move and that any part of the icon can be selected for the right information. B is distracting and falls prey to the detriment of the others which is the lack of control of "information presentation speed"
	I liked when only the needed information was shown in a non-distracting manner. A took too long and blocked the other moving objects. I don't remember B happening.
	I find A more preferable due to being able to compare to the query faster than the others.
	D/C scrolled too fast for locations; B flashed too fast; A wish I could click and only display one whole line
	In concept, D there wasn't enough time to see the information before it went away. Concept A was preferable because the info was already there all the time so all I had to do was fin d the correct line of information to see it.
	"A" lets you see all the information at once, which can be quicker. The rest of the options tend to be slow and risk missing information.
	"A" was the easiest and most convenient. Seeing all info up front is better than scrolling or flashing text. Concepts B-D create more clicking and add more pressure.
	Scrolling left to right was difficult while keeping attention on other tasks.

Y/N? Did you find the function buttons around the entity tag to be helpful?	**If yes, how and why?**
Key: A = Pop up Dialog Box B = RSVP	Yes, with all options except the pop-up box. I would often miss a piece of information and have to click on the corresponding button to look again.

C = Times square D = Cinema Credits	
	I used the rewind buttons pretty often. That was it but it made things simpler.
	When the text was scrolling or flashing, I could click on the boxes to bring it back to the number query I was reporting on.
	PID - thought the letters were good for what they meant. Maybe could have been separated a little more. Maybe a little "dead zone" between them would help with "mousing/clicking" errors.
	Yes, they made it very easy to get exactly where I wanted to be. I almost always used them to view information.
	I would keep passing the button over and over to refresh so I could double check.
	I would occasionally need to rewind the line, especially for the coordinate lines.
	I used the run/pause/stop functions far more frequently than others. Generally, if I didn't see something the first time I would re-select the line function which brought me directly back to the beginning.
	Allowed me to quickly jump to the right function instead of having to scroll through.
	Yes, but the added step of finding which button is a number completely absorbs attention for a moment even though it is much better than waiting. Ideally, each button would be associated to an explicitly unique object such that your brain doesn't require sorting/processing.
	Sorted the tasks. Didn't overwhelm the screen with jobs to do
	They were useful to go back. I had to use them for concept ID C&D
	It was nice to be able to click the function buttons again to start at the correct line so I could see the info again if it was a scrolling concept ID.
	If I missed any info, I could just click back to it
	Useful to find specific information.

Y/N? Did you find the function buttons around the entity tag to be helpful?	**If no, did you use them? Y/N**	**Participant Comments:**
Key: A = Pop up Dialog Box B = RSVP C = Times square D = Cinema Credits		
	Y	They serve only to select the right line, and in some cases of overlap could not be properly selected.
		I didn't use the arrows on the side. I just re-clicked the box I needed to see the message again.
		No. I tried using it once or twice but it felt less natural and higher potential for error.
	N	No, No, if you are talking about the rewind and fast forward buttons.
	N	I used the buttons except fast forward and rewind. I would click the same button twice to repeat the answer instead of moving my mouse to click fast forward or rewind.
	Y	It was too much to have to read quickly (while keeping an eye on the other task) and having to figure out how to rewind/fast forward/etc. only helpful to rewind if text went too fast.
	N	
	Y	
	N	
	N	

Y/N? Specifically regarding the function buttons around the entity tag, were the "fast forward" or "rewind" buttons helpful?	**If No, did you use them at all? Y/N?**	**If yes to either, please indicate why and how.**

N	Y	I only ended up using the rewind button or fast forward once or twice. I usually just clicked on the ones around the entity.
Y		Like I said in question 3, I used rewind pretty often. I feel like the task would have been much more painful without it.
Y & N		Yes because you could move to the previous or next query information quickly. No because sometimes it was scrolling so fast or I had missed it and it didn't go far enough back or forward. So it was sometimes just easier to use the function buttons.
N	N	Did not use cause it was easier to click on the line box even though not labeled.
N	N	
N	N	
N	N	
Y		See Q #3 (I would occasionally need to rewind the line, especially for the coordinate lines.)
N	N	Did not use, sometimes reselected the line button to reset.
N	N	
N	N	
N	N	
		I never used them, but this again feels like a constant brain usage task that would increase probability of task failure. Why not have a pause button?
N	N	
N	N	
N	N	
N	Y	Only helpful to rewind. Accidentally fast forwarded a few times. Rewind was rarely needed.

Y		I never fast forwarded but I would rewind
N	N	
N	N	
N	N	
Y		I fast forwarded on the scrolling left/right presentation style because it would take too long to get desired info if I didn't.

Y/N? In some cases, while answering the queries, the text tag remained fixed in place on the screen while other times it continued to move with the entity. Do you think this affected your task performance in any way?	**If yes, please indicate why and how?**
Key: A = Pop up Dialog Box B = RSVP C = Times square D = Cinema Credits	Yes, I did not like that I had to click twice to get the tag to follow the entity. This often distracted me and diverted my attention.
	I feel that when it stayed in place, my concentration was broken because the other way (constant motion) felt more natural to me
	When it remained fixed, I could perform the query task more efficiently. When it continued to move without the entity, I got distracted wondering if it's changing directions or color. I know the query task was most important, but the secondary task started to be more of a concern at these times.
	Was not distracted by the tag moving - too much going on with the tag already
	Yes. I often diverted my attention to clicking the box again so it would reattach to the entity.
	It was easier for me to read when the tag was stationary.
	Yes, because I felt more rushed to answer instead of losing the entity from the screen
	In cases of "A" this is less important due to the amount of time left up on screen, but in other cases, the moving box is preferable as it makes tracking multiple objects for directional changes easier.

	Yes, because I would forget to click it again to reattach it. It was distracting when you are already multi-tasking and seemed useless.
	Yes. It was weird to have the text stay in place because then you were trying to read and track the object.
	Yes, it requires more focus to track an object that is moving.
	As my concentration on the task began to waiver I found myself having to click again to reread.
	If the text stayed in place, I found it would get in the way of watching the entities. It was also another task to have to click on it to get it to move again which drew my attention away from the main tasks.
	I think it affected my performance because it might have been easier or harder to read text boxes that moved compared to holding still.
	When text was on screen and stayed there it was easier to find what was needed and click answer quickly without having to scroll back and read again.
	It was a little distracting at first but after a few times, I got used to it.
	It is easier to watch a non-moving screen as opposed to a dynamic one.

Please indicate below any general comments you have regarding your experience with this study and/or suggestions for improvements for the text presentation concepts.
Key: A = Pop up Dialog Box B = RSVP C = Times square D = Cinema Credits
In general, I really liked this task! Because I am biased towards liking the pop-up box the best, my main suggestion is (if you want us to do better on the secondary task) to make the box transparent or somehow make it to where you can still see the entity that hides behind it. Regarding the scrolling or flashing text boxes, possibly slow them down some. Regarding the query --> because there are many numbers in the task, it would be easier to recognize the query number and line number faster if it was bolded. Example: Entity 1 Line 5 xx xxx. I'm not sure how often this happened, but I caught myself switching these around occasionally. Thanks!
Possibly label lines. Take off rewind and fast forward - this could allow to expand the tag to show the line boxes better and the text inside, and to make room to put numbers in line boxes. Would PID boxes

set up vertically or in a circle (might argue that this has "more spatial" distinction) be better than horizontal as in the task?
Better confirmation feedback when making a selection. For example when selecting P for present, the P flashes green quickly then disappears. Separate the P.I.D buttons to prevent misclicks. Mouse speed and lack of mousepad led to errors. Suggested having query response (Y/N) at the top next to the question instead of by the entity.
Issue with an earlier trial (#5-8?) where one popped up mid-trial saying "GUI MFC application stopped working." The trial continued running in the background and I closed out of the error using the red "x" on the error box and resumed the trial. This happened again in trial 22 and I repeated trial 22.
I think concept B moved a little too quickly through the lines.
The predictability in this experiment is high due to the use of the identical sequence of movement, identity changes and presentations even down to timing for queries in text making it both repetitive and takes away from the overall difficulty and interest in the task. One might use multiple, staggered sequences to combat this.
Combine concept A and the flashing, by enabling a selection of the line and then the popup only displays the selected line. This would take up less real estate and be quicker than any of the other flashing concepts, especially for the loc queries. I often had to reselect the number twice, because I couldn't compare the coordinates fast enough while it was scrolling.
The only thing I really noticed, not sure if it was intentional or not, was that every test would be the exact same with the exception that the entities would switch directions from where they appeared to go.
When reading sequences of numbers with over 4 characters the text needs to move slower. Reading heading north and comparing is way easier than reading target 10C 36.19/95.88 in the same time frame.
As someone who has played very intensive multi-tasking games, to optimize information processing you need to find/create cues that require minimal pursing per task. Such as finding line N, the input of calculating line n to which block + constant mouse/motor control + delay time in displaying correct information is very cumbersome. The large text box makes this task much easier by allowing higher variance input, not calculating, and no delay time. As an extended example if display time of required text is randomized, multitask performance will decrease because of overloaded visual information. If the text box was always the same spot on the side of the side of the screen, contained all information in the same spot, and flashed to show you successfully clicked a new entity there would be minimal over digestion. Also, for selecting present, identity or direction the box click precision was a major cause for errors. Maybe instead of flash when selecting a box, highlight it, and use error key based input. This not only reduces task load but separates the motor skill of each task.
The task was very boring. It felt like my performance decreased even as I got more practice. Perhaps a punishment for making mistakes, like a shock collar, would have helped me keep better attention.
I don't like how the options went through the whole list of information, which made me have to click it again to stop the scrolling. I'd rather it repeat the information I need over and over, instead of moving onto the next line of information. I think a good way to improve the experiment is by adding the keyboard to answer the query. For example "Y" for yes and "N" for no, instead of having to click with the mouse. Maybe even being able to press 1, 2, 3, etc. instead of trying to click the little tabs around the box to get the line of information you are looking for.

At times I wasn't sure if I actually clicked the button I needed to on the correct object. Maybe if the P/I/D or Y/N flashed or made a sound or something to ensure that the subject knows what was clicked. That is better than clicking on it twice to make sure or just not knowing if it was done correctly.
I would suggest having set break times.
After about 5 blocks I could recognize the patterns of the entities so I think that made it easier than it was at the beginning.
All my thoughts are pretty much covered in question 1. It was an interesting task.

Appendix D: Institutional Review Board Protocol Document

A Human Performance Comparison of Text Presentation Techniques During Information Extraction Using A Dual-Task Paradigm—Phase 1: Secondary Task Difficulty Parameters Determination and Phase 2: Human Performance Comparison Study.

1. **Principal Investigators**
   Eric E. Geiselman/DR-III/Senior Engineering Research Psychologist
   711 HPW/RHCV

2. **Associate Investigators**
   Michael E. Miller/AD-23/Associate Professor
   AFIT/ENV

   Eric Heft/DR-II/Computer Engineer
   711 HPW/RHCV

   Lauren E. Menke/Ctr
   711 HPW/RHCV/Ball Aerospace

3. **Medical Consultant or Monitor**
   Anticipate N/A: due minimal risk determination based on the benign and familiar nature of the experimental apparatus and data collection location.

4. **Facility/Contractor**
   The testing facility for this study will be the RHCV Visualization Laboratory Room 305 or 306, Bldg 248, 2255 H Street, WPAFB, OH. The laboratory contains equipment utilized for human visualization performance experimentation.

5. **Objective**
   The phase 1 study was conducted in order to determine the appropriate level of difficulty for a multiple entity secondary task being developed as part of series of studies planned to investigate the utility of different text presentation techniques. The phase 2 text presentation study will use a secondary task paradigm for objective performance measurement. Manipulation of the text presentation techniques will form the basis of a primary task. The present study will be performed to determine which text-based presentation and access techniques best aid human performance.

6. **Background**
   The objective of information visualization is to afford an observer (user) the appreciation of complex data relationships in way that is easy to understand via the visual modality. Among the many reasons for visualization use, a high-level

distinction can be made between information discovery and information interpretation. Certain visualization formats lend themselves more toward one objective versus the other. Simply stated, visualization format choice and its associated information extraction value is task dependent. From the user perspective, one point of commonality across the visualization spectrum is this: format usefulness depends on how well the information presented supports the process of effective decision making. Information accessibility, interactivity, and efficiency are also desirable common features across a wide variety of visualization applications. To realize effective visualization, information accessibility, and interactivity, it is not uncommon that graphical data visualization presentation be augmented with alphanumeric (text) symbols for the purpose of labeling, presenting detail, defining specific values, etc. In terms of efficiency, the need to incorporate text into a complex visualization can be very costly. This is especially true considering the display surface area or "real estate" required to ensure that visually displayed text be readable. Well established standards are in place to define the physical dimensions required to support symbol readability and text modifier use (MIL-STD-2525 Common Warfighting Symbology; AC 25-11 Electronic Flight Deck Displays). Either on the written page or via an electronic display, minimum levels of brightness, contrast, character size, and spacing must be achieved to insure acceptable readability (Kruk & Muter, 1984). The objective of visualization can be summed up by the classic axiom: "a picture is worth a thousand words," but what must a word be worth to justify its inclusion in a picture?

In many applications, there is often more information the interface designer wishes to make accessible to the user than there is physical display surface area available. Regardless of the display scale under consideration, be it a data wall or a cellular phone screen, when text presentation is deemed necessary, the efficiency with which it can be included in the visualization is challenging. If a lot of text is displayed at once, such as inside a pop-up window, an occlusion and clutter cost is incurred. When attempting to display text within a small area, a readability and/or accessibility cost is incurred. For this effort, the concepts of interest share a simple goal: how can the amount of text-based information available to a visualization user be maximized while minimizing the display surface area requirement for its availability? The researchers at 711 HPW/RHCV are studying human performance during information extraction tasks for purposes of determining which text presentation techniques offer the best performance. To quantify performance, a dual task paradigm will be utilized. The present study is being performed to objectively determine the difficulty level of the secondary task and insure appropriate sensitivity of the metric.

The secondary task, also referred to as the "global" task is designed to mimic a generic scenario monitoring task which may be performed by operators such as air traffic controllers or battle management analysts. The task requires a participant to monitor the activity of several entities and report observed entity status changes when they occur.

7. Impact

The present study will determine the appropriate level of difficulty of a task set intended to form the basis of the global (secondary) task for purposes of measuring performance of the local (primary) task in an experiment which will be performed in the future. The phase 1 and 2 series of experiments will be performed to compare various text presentation techniques on their information conveyance merit during the performance of complex monitoring tasks representative of command and control. We anticipate that the findings of these studies will generalize to the ability to predict text presentation performance across a wide variety of applications and help populate the scientific knowledgebase in support of large, complex data set visualization, and decision support system interface design.

8. Phase 1 Experimental Methodology

Apparatus:

1) Display source: A standard LCD monitor will be utilized to present the visual stimulus to the subjects. The displayed information is scalable so that multiple monitor sizes and aspect ratios can be used as the display source without affecting the interface appearance. The minimum display size used for the current study will be 17” measured diagonally. Subjects will be seated at a normal distance from the display source representative of the viewing distance to a conventional workstation to insure a similar comfort level.

2) Computer: A conventional computer equipped with standard keyboard and input devices will be used to generate the stimulus display and for subject response input/questionnaire responses. Questionnaire responses will also be collected via pen and paper recording.

3) Software: Code development for this study was performed in-house by associate investigator Eric Heft. Software development includes generation of all the visual elements, task mechanization, performance data collection routines, and raw data recording.

Subjects:

A maximum of ten (10) subjects will perform the current experiment. The number of subjects may be fewer based on how consistent performance is from subject to subject. From past experience, it is anticipated that data from at least 5 subjects will be required. Each subject will participate in all conditions of the study. Subjects will be males and/or females between the ages of 18 to 60 with self-reported normal or corrected to normal visual acuity and color vision. Subjects will be recruited among local area personnel (active duty, civilian and contractor). Subject recruitment will be done via “word-of-mouth” and/or email by an experimenter. The following is an example of the phrasing to be used for the verbal recruitment:

> “We are looking for people to participate in a study examining the work load and difficulty of a tracking task including the collection of performance measures during interaction with visual events. In the study, you will be asked to watch for specific events occurring on a workstation display while reporting the occurrence of specified events via a mouse input device. The task will become increasingly more difficult over the course of each session. We are interested in knowing when performance begins to deteriorate depending on task difficulty. The time required to complete the study is no more than 2 hours and takes place in the Visualization Laboratory in building 248. Would you be interested in participating?”

Those expressing interest will receive an email containing a detailed description of the study, an outline of the minimum vision performance requirement, and the expected duration required for participation. An example recruitment email is attached to this document. Handedness will be recorded via a self-report questionnaire but will not be controlled as a potential systematic cause of variability. Subjects are not expected to benefit directly or be compensated monetarily or otherwise, for their participation in this research study.

Duration:

Approximately 3 months will be required to complete data collection for the experiment. Each subject’s total participation duration including introduction/consent, instructions, demographics/vision screening, data-collection, and rest breaks will be less than 120 minutes.

Questionnaires:

1) Demographic Questionnaire (see attachment): A questionnaire, consisting of background information questions (e.g. age, vision correction, handedness, and video-gaming experience) will be administered prior to data collection trials. Subjects will input their responses using conventional pen and paper response input. The questionnaire responses may assist during interpretation of the findings if variability was found to be systematically influenced by physiological or experiential factors.

2) Workload assessment: Workload reporting will be performed by having subjects complete a Bedford Workload Scale (Roscoe and Ellis, 1990), decision tree procedure. The Bedford Workload Scale rating scale for self-assessed workload is a fairly self-explanatory procedure where subjects report a subjective workload rating based on consideration of spare capacity to perform additional tasks with regard to the presently performed tasks. Use of the Bedford Workload Scale will be described during subject introduction and instructions. A practice data collection (task training) session will include demonstrating use of the scale. Figure 1 shows the Bedford Workload Scale decision tree.

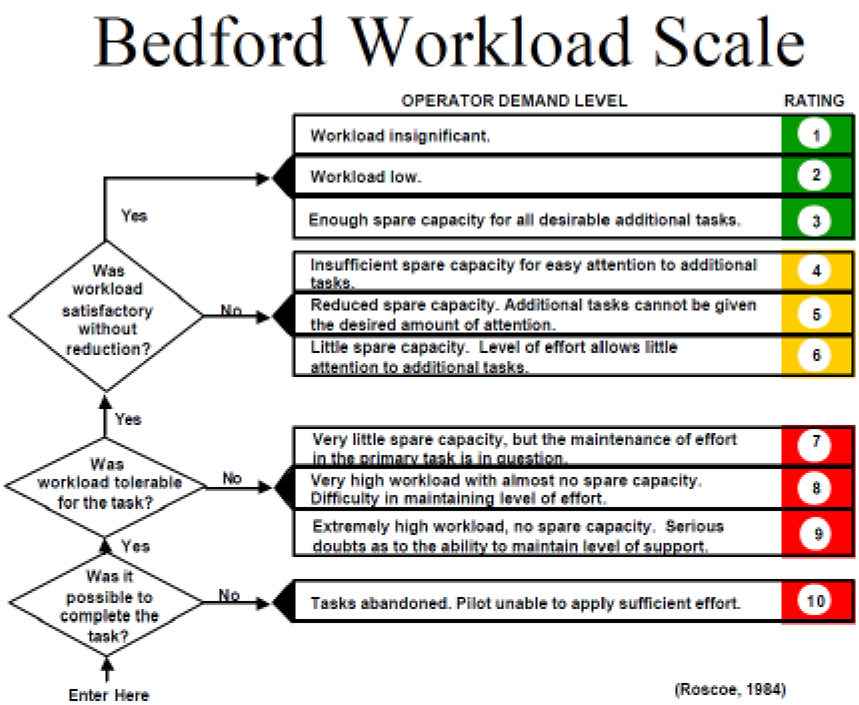


Figure 1. Bedford Workload Scale graphical decision tree.

Data Collection Procedure and Sequence:

Subjects will be seated at a table equipped with desktop monitor, keyboard, and mouse. The experimental apparatus will mimic a conventional computer workstation. The experiment will consist of the following sequence: 1) introduction/consent, 2) demographic questionnaire, 3) task training, and 4) data collection.

1) Introduction/safety briefing/consent: Subjects will read a short description of the study rationale and sequence of events. They will then receive a short safety briefing explaining what to do in the event of an

emergency. This briefing will include where the exits are located, tornado and shelter-in-place locations, and where to assemble if the building is evacuated. Subjects will read the consent form and sign it to confirm their agreement to participate in the study. Subjects will be informed that their participation in the study is completely voluntary but dependent on their reporting normal or corrected-to-normal vision. Subjects will be informed that they may choose at any time throughout the experiment to terminate their participation. A copy of the consent form is attached below.

2) Demographic questionnaire: Subjects will complete a questionnaire to report past experience with video displays and tasks which are similar to those used in the study.

3) Vision screening: Subjects will be asked if they have normal or corrected-to-normal visual acuity and normal color vision. Corrective lens should be used for any and all phases of participation in the study if worn normally for similar computer tasks.

4) Task training: subjects will be given detailed instructions regarding the task operation, objectives, and desired performance. Subjects will also be given practice during a fully-dynamic task session. The practice session will include specific instruction for, and hands-on experience with all interactions they are asked to perform during the actual experimental events. The entire range of input and output variations will be covered during this training period. It is anticipated that this training will last fewer than 30 minutes. During the training session, subjects will be able to stop and ask questions concerning the tasks. During this time, subjects will practice completing the Bedford Workload Scale self-assessment procedure. Once the experimenter and subject are satisfied that the tasks are understood; adequate performance has been demonstrated, the data collection session will be initiated.

5) Data collection: Three data collection sessions will be conducted for each participant in the study. The number of trials within a session will depend on how quickly each participant reaches the predetermined performance criterion. Each session will consist of continuous performance of trials for a duration of fewer than 25 minutes each. A rest break (a duration of approximately three minutes) will be scheduled between running the first/second and the second/third data collection sessions. Subjects will be informed via the consent procedure that they may take breaks or discontinue their participation in the study at any time. In the event that a session is interrupted during a normal run, the event will be noted and recorded. Similar documentation will be generated for

excessively long break periods, voluntary discontinuance, and/or equipment failure related interruptions.

6) Workload assessment: Workload reporting will be performed by having subjects complete a Bedford Workload Scale decision tree procedure at the completion of each discrete trial within the sessions. The Bedford Scale is a rating scale for self-assessed workload is a fairly self-explanatory procedure where subjects report a subjective workload rating based on consideration of spare capacity to perform additional tasks with regard to the presently performed tasks.

Task Description:

The task is designed to mimic a generic scenario monitoring task which may be performed by operators such as air traffic controllers or battle management analysts. The task requires a participant to monitor the activity of several entities and report observed entity status changes when they occur. Reportable status changes include changes of entity direction, entity identification, and when an entity enters into the display area (called presence). Figure 2 is a screen capture depicting the presence of a number of entities in the scenario. The entities in the scenario move across the screen at a constant rate in one of four possible directions relative to screen coordinates: up (+Y), down (-Y), right (+X), or left (-X). Entity identification is differentiated via fairly standard shape and color coded symbols: Friendly (green circle), Hostile (red diamond), and Unknown (blue square). Attached to each symbol is a data tag intended to provide specific identification labeling and other status information regarding the associated entity. This data tag is a component of the primary (or “local”) task to be implemented for a follow-on study and will not be discussed in detail here. For this study, the data tags will be included within the global task development study to more closely represent display appearance of the full-task study. It is possible that the data tags cause some distraction or occlusion effect. Including the data tags was maintained to control for any possible confounds that may otherwise be caused by their absence in this study and their presence in follow-on studies.

Entrance of new entities (players) into the scenario will be reported as presence. According to a controlled (manipulated) schedule, players will enter and exit the scenario. Once an entity enters onto the screen, participants will be required to report that the player is present. All present entities will then be monitored for identification and direction changes while they are visible on the screen and these occurrences will be reported as well. Performance degradation will be determined based upon the number of times the participant misses or fails to accurately report these events.

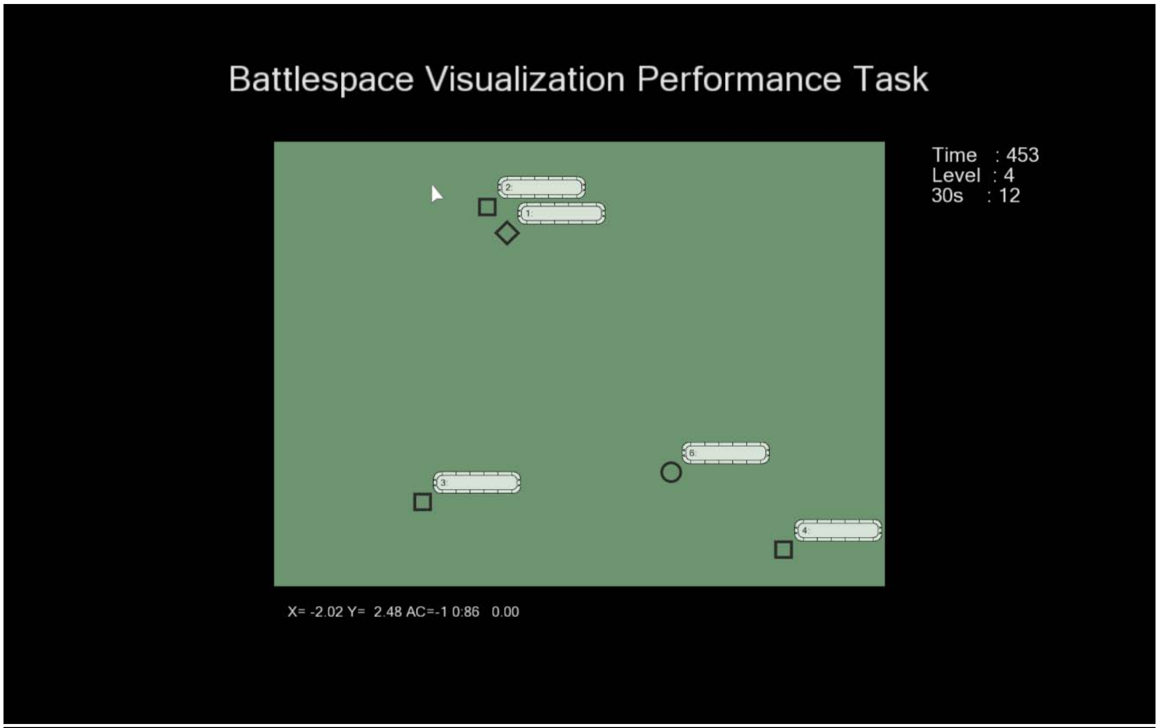


Figure 2. Screen capture with all three entity types.

Participant Interaction:

All direct interaction between the participant and the scenario is accomplished with a standard "mouse" input device. The mouse is used to move a cursor to any location within the monitor (display) screen area. Proximity of the cursor to an entity automatically associates that entity to the cursor for purposes of selection and further interaction. A right "click" of the mouse device button is used to select the entity for entity change reporting. Once an entity is selected with a right click, a small dialog box is presented so the nature of the entity change can be logged. Figure 3 shows a selected entity with the input box comprised of three soft buttons. Hovering the mouse cursor over one of the three soft buttons gives feedback in the form of a highlighted soft button. The buttons are labeled with a "P" for reporting presence of a new entity, an "I" to report an identification change, and a "D" for direction change reporting. Once a player is selected and the reporting buttons are present, the reporting button dialog pops up and remains stationary relative to screen coordinates (fixed in position) and player movement until a second left click is input by the participant on the P, I, or D buttons within the "PID box". Figure 3 depicts a static PID box that is associated with "Player 2" even though Player 2 has continued to move away from the location where it was selected. The location of the entity selection is approximately where the PID box is drawn. The PID box will remain visible until a P, I, or D button is selected via a left mouse click (PID choice condition), a right click is made close to another player or screen area (cancelation condition), or 10 seconds has elapsed without any input made by the participant (timeout condition). Only one PID box will be present at any one time during the trial. Once an input is registered within the PID box (PID choice condition), that PID box is removed and will not be present again until a right click is made on any player in the scenario.

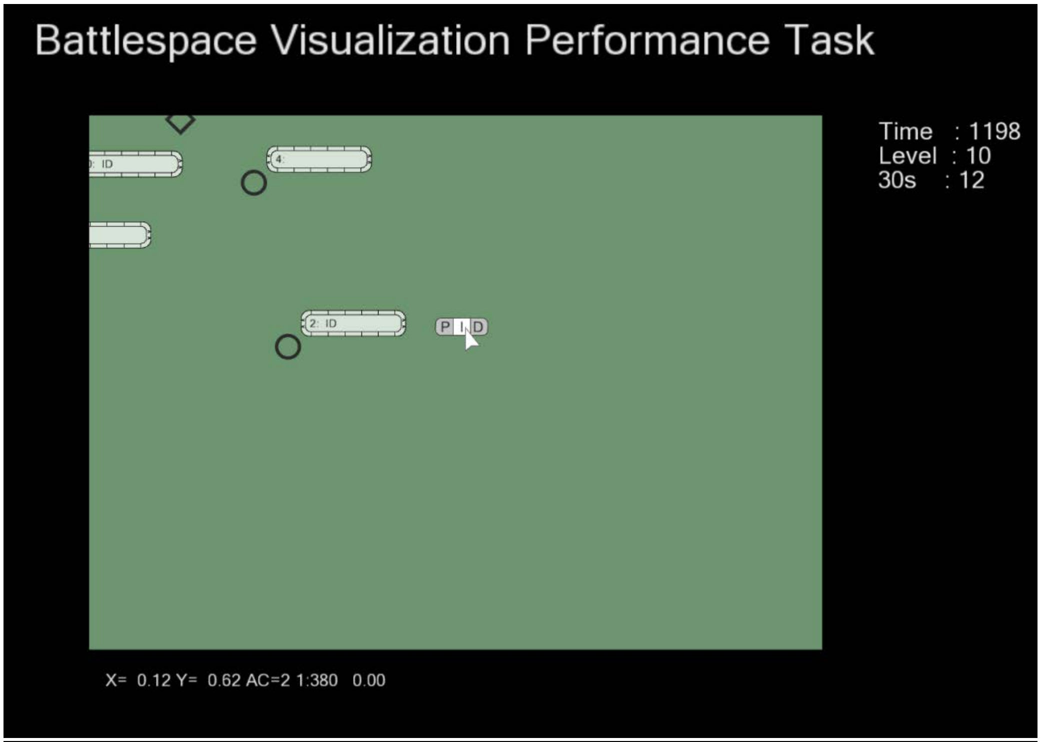


Figure 3. Screen capture with “PID” selection box and “identification” highlighted.

Task Difficulty Manipulation:

The objective of the present study is to determine the appropriate level of performance difficulty for a secondary task that will be used as a global task for a future full-task scenario where the independent variables of interest will be evaluated. For the secondary task, it is desired that the variation presented to the participants meets the following criteria: 1) the task should be challenging (but identified by the participant as acceptable workload); 2) the task should be possible to complete with close to error-free performance; 3) the events of the task should not be predictable by the participant; 4) the variation of the events experienced by participants is minimized between participants. These are the elements of an effective and sensitive secondary task with minimal systematic confounding variation threat to the dual-task paradigm experiment. For the dual-task approach, performance differences within the secondary task will be attributed to the manipulation of the primary-task independent variables of interest.

Secondary task difficulty will be manipulated by using a duration-based ramp-up technique. Difficulty will be incrementally increased until the participant fails to maintain minimum criterion performance over a specified period of time. Task difficulty will be metered by increasing the number of players entered into the scenario, by increasing the entity identification change frequency, and increasing the frequency of entity direction changes. Since each of these events require input interaction by the participant, it should follow that the task will become increasingly busy and thus, increasingly difficult.

Selection of event frequency is manipulated by incrementally decreasing the start and end duration between events (moving time window). For player entry, the decision is a simple yes or no determination. The reason for having a minimum start time is to avoid cases where events may occur simultaneously within a single category. The maximum time value is used to define the decision frequency and related frequency of event changes. For the "ID" category the decision options include not changing (null change) or changing to one of the two remaining ID types. A similar time window range is used for the direction change variable. The decision applied to the entity for this category includes a null change option or a direction change + or – 90 degrees from the present direction of travel. If a direction change occurs, it is applied in total with the next simulation frame update (the direction change is not smoothed in). Levels of performance are determined by amount of time allotted within the minimum and maximum duration intervals. Each duration change is a discrete "level" of scenario difficulty. Within a difficulty level, participants will perform for a maximum duration of two minutes before advancing to the next level. Performance feedback in the form of the achieved level and number of errors within the preceding level will be displayed for the participants. Required performance criterion presently allows participants to make a maximum number of errors of any type within a specified moving time window to continue advancing levels. If the error criterion is exceeded, the trial automatically ends and the highest performance level achieved by the participant is recorded. The moving time window allows 3 errors to occur within a primary 30 second period. When this happens, and 3 more errors occur within a subsequent 30 second period, the exit criteria is triggered and that trial ends. The performance criterion threshold as well as the attributes defining a level can be adjusted within the simulation code.

Task Difficulty Determination:

The appropriate level of difficulty for the secondary task will be determined by this study. Participants will perform the task several times to define the highest level of incremental task difficulty that can be consistently achieved under reported conditions of acceptable workload (not working too hard and not too easy while maintaining very good accuracy performance). An accurate input is defined by the selection of the appropriate entity and PID box button for the preceding change event. Errors are defined several ways: 1) failing to select an entity within 10 seconds of a PID event; 2) failing to select the correct PID event change button within 10 seconds of player selection (PID button selection must be made within 10 seconds of entity selection or it will be reset); 3) selection of the appropriate entity but selecting the incorrect PID event button; 3) selection of an entity not associated with a preceding change.

Using this ramp-up difficulty technique, each level of the scenario is defined by frequency of events occurring and the approximate number of players in the scenario. To determine the overall task difficulty achieved across all participants and trials, the highest difficulty level successfully achieved by each participant will be recorded for analysis. The mean of the summed highest level will be used to define the change attributes for the secondary task. For instance, if the mean highest achievable level is 10, given the present ramp-up schedule, an add player decision will be made at a rate defined by 11.97s(min) and 23.95s(max) run time. An ID change decision will be made for each player at a rate defined by 11.97s(min) and 17.96s(max) run time. Similarly, a direction change decision will be made for each player at a rate defined by 11.97s(min) and 17.96s(max) run time.

Task Scripting Technique:

Once the parameters which define the appropriate secondary task difficulty (min/max durations for player entrance, ID change, and direction change, the trial duration required to reach mean criterion performance will be used as the basis for producing an objectively derived secondary task scenario for the follow-on study. A hardcoded script of the events defining the secondary task will prescribe the events to occur within the scenario so that each participant for each trial will experience the same overall number of events occurring with the same frequencies of occurrence. To help ensure that the scripted scenario is not predicable and potentially learned, a technique will be employed to use reversals of the events such as identity and direction changes. This allows all participants to experience the same overall global task while reducing unwanted variability across the statistical model.

Experimental Design, Data, and Data Analysis:

Because this is not a comparison study, experimental control for purposes of statistical analysis is not necessary.

Task accuracy and event reaction time will be collected to plot objective performance over time. Overall duration (based on run time) of each trial will be recorded. The performance data will be output to a spreadsheet for analysis. After each trial, participants will indicate a workload rating for the trial based on a subjective feedback rating. The purpose of the workload rating is to determine that task workload is acceptable. Since this study is being performed to determine the appropriate difficulty level of a secondary task, it is important that acceptable performance is achievable but there should be some workload capacity remaining as a resource for the eventual performance of the primary task.

9. **Phase 2 Experimental Methodology**

Apparatus:

1) Display source: A standard LCD monitor will be utilized to present the visual stimulus to the subjects. The minimum display size used for the current study will be 17" measured diagonally. Subjects will be seated at a normal distance (approximately 60 cm) from the display source. This is representative of the viewing distance to a conventional workstation and associated level of comfort expected of a seated worker.

2) Computer: A conventional computer equipped with standard keyboard and cursor input device (mouse) will be used to generate the stimulus display and used for subject response input/questionnaire responses. Questionnaire responses will also be collected via pen and paper recording.

3) Software: Code development for this study was performed in-house by Eric Heft. Software development includes generation of all the visual elements, task mechanization, performance data collection routines, raw data recording, and any automatic data reduction.

Subjects:

Twenty four (24) subjects will participate in the proposed experiment. The number of subjects was determined by comparison to similar studies where adequate statistical power was demonstrated for the research purposes (Broadbent & Broadbent, 1987) as well as 24 subjects is the minimum number required for a fully counterbalanced experimental design. Each subject will participate in all conditions of the study. Subjects will be males and/or females between the ages of 18 to 60. Subjects will be recruited among local area personnel (active duty, civilian, and contractor). Subject recruitment will be done via "word-of-mouth" and/or email by an experimenter. No specific skillset will be required. Subjects will be required to self-report normal or corrected-to-normal visual acuity and color vision.

Demographic Questionnaire:

A questionnaire, consisting of background information questions (e.g. age, vision correction, handedness, computer usage, and video-gaming experience) will be administered prior to data collection trials (See Appendix A--Demographic Questionnaire section). Subjects will input their responses using conventional pen and paper response input. The questionnaire responses may assist during interpretation of the findings if variability was found to be systematically influenced by physiological or experiential factors.

Duration:

Each subject's total participation duration including introduction/consent, instructions, demographics/vision screening, data-collection, and rest breaks will be less than 320 minutes.

Text Access Technique Development Description:

This discussion will introduce the various techniques and user interaction concepts designed to provide access to a relatively large amount of readable text in a small display area. The display area of interest will be constrained roughly to that which one would expect to be used for a label of an entity within a C^3I data presentation with the exception that display real estate is taken up by the control mechanization graphics added to non-baseline designs. As such, a logical choice for the text content of interest is the type of information typically found within a standard 9-line information set used to pass targeting information between command functions and mission performers.

Example Nine (9)–Line Used in the Context of Close Air Support (CAS):

The objective of the 9-line communication protocol is the presentation of vital data elements to, in this case, support a targeting or sensor data collection activity. A top level requirement for this communication sequence is the identification of the acting agent or performer of the mission element. For C^3I display purposes, this item is typically a call sign label. The 9-line information underlying that call sign allows command and control functions to understand the assignment and intent of the acting agent (Air Land Sea Application Center, 1997). The contents of the 9-line can be thought of as a "briefing" of the critical information items required to perform a mission assignment. For clarity and ease of use, the items contained in the 9-line are formatted in a standard sequence. In support of decision making activities, access to the 9-line affords monitoring and progress situation awareness at the single entity level. Also, as mission changes occur and new assignments are passed to entities of interest, 9-line information element availability forms the basis of a closed communication loop between command functions and actors. The elements of the standard 9-line sequence are as follows within Table 1.

Table 1. Contents and descriptions of 9-line elements.

Identification (ID) tag: call sign label
1) Initial Point (IP) or Battle Position (BP): Ingress point expressed as a landmark, waypoint, latitude/longitude coordinates (lat/long), etc.
2) Heading: From IP/PB to target expressed in magnetic compass coordinates.

3) Distance: From IP/BP to target expressed in nautical miles (from IP) or meters (from BP).
4) Target Elevation: Expressed in feet/mean sea level (MSL).
5) Target Description: As appropriate.
6) Target Location: Visual, lat/long, grid, or offsets.
7) Type Mark: Laser, waypoint, infrared, or beacon. Code: Actual code.
8) Location of Friendly Forces: As appropriate. Position Marked By: As appropriate
9) Egress: As appropriate. Remarks: As appropriate. Time On Target: Expressed in Universal Time Constant (UTC).
Remarks: Additional information as needed.

Baseline Display Condition (Pop-Up Dialog Box (DB) Format):

For purposes of comparison, a baseline display technique, or format, will consist of a conventional pop-up dialog box approach for text access. Each active entity in the aggregate display will have an associated ID data tag attached to it. The data tag box will be a fixed size able to accommodate approximately 13 characters. The tag box (ID tag) size is based on the human ability to perceive 12-14 characters in a single fixation with acceptable accuracy (Just & Carpenter, 1998; Robeck & Wallace, 1990; Rayner, 1998; Öquist, 2006). Figure 4 (as presented by Öquist, 2006) demonstrates the foveal and parafoveal accuracy approximation within a single human eye fixation.

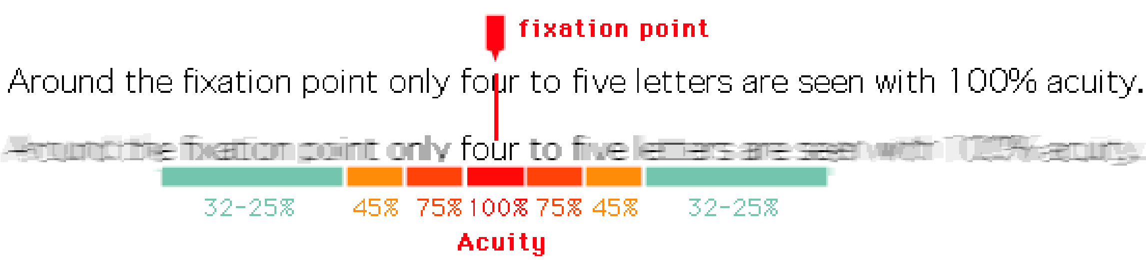


Figure 4. Illustration of single fixation foveal and parafoveal perceptual accuracy.

For this investigation, it is desirable that the entity label be perceived with a single fixation and the alphanumeric characters of the ID tag should require little to no saccadic eye movement to be read in their entirety (Figure 5). For access to the underlying 9-line information associated with the entity, an observer will select the ID tag via a conventional pointing device (hovering a mouse-driven cursor over the ID tag and actuating with a left-click). This action will result in the display of a pop-up dialog box containing the text of the completed 9-line information for the entity of interest. The vertical size of the 9-line dialog box will vary depending on the amount of text required of the 9-line assignment. Text will be wrapped to fit within a window size of defined by the often used “page” format 50 character-wide limitation (Piolat, Roussey, & Thunin, 1997). Figure 6 shows a rendering of a completed 9-line page pop-up dialog box. The window will stay open until the cursor is within the window and performs another left-click via the mouse controller. This interaction feature enables the user to keep multiple windows open for multiple entities as long as desired. It is expected that this interaction mechanization will be familiar and intuitive for users and will require little or no usability training.

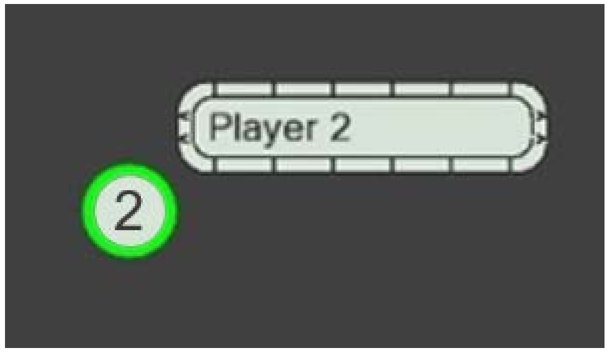


Figure 5. Entity with identification data tag label.

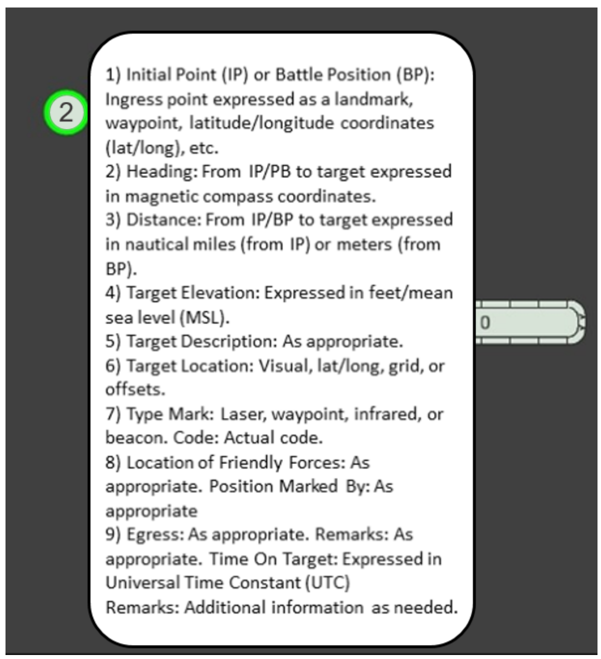


Figure 6. Example entity with pop-up dialog box containing 9-line information.

Spatially-Constrained Concepts:

Figure 7 shows a rendering of a proposed spatially-constrained entity tag with user interaction symbology. Similar to the baseline condition, the drawing area of the data tag will accommodate approximately 13 characters at any one time. This will be the maximum size of the drawing area regardless of the mechanization of the underlying text presentation variations (with the exception of DB baseline shown in Figure 6). It is likely a requirement that some ID label stay with each entity simultaneously with the cycled 9-line information within the data tag drawing area. This will be accomplished by superimposing an ID number within each entity shape (Figure 8).

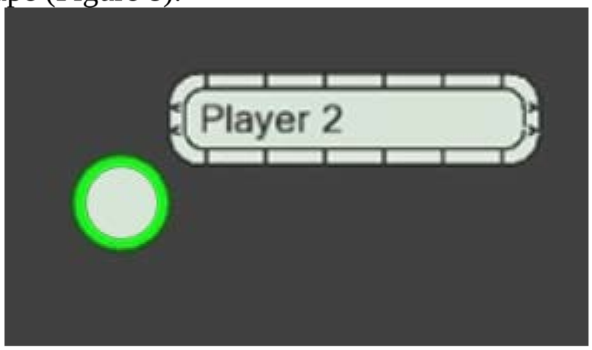


Figure 7. Example spatially-constrained presentation window and function keys.

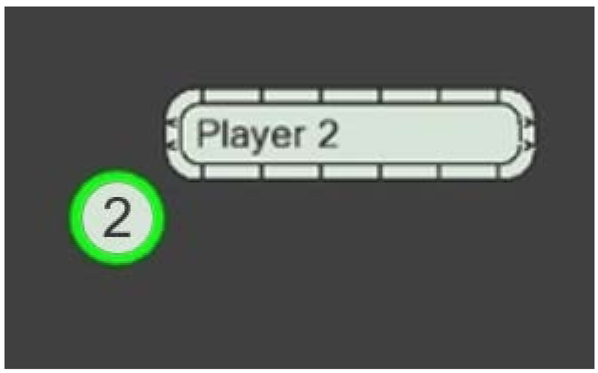


Figure 8. Entity shape with the addition of an ID number.

When selected by the subject, the 9-line information associated with the entity of interest will become active and the presentation will continue until the entity is “de-selected” by the user. This functionality will allow the information associated with multiple entities to be available simultaneously. Text presentation techniques to be considered include: RSVP, TS, and CC. For purposes of evaluation, the DB format represents the non-interactive baseline text access technique.

Interaction and Feedback Concepts:

The spatially-constrained text access formats will require significantly more interaction between the user and the presentation formats compared to the DB technique. This is especially true where information seeking tasks require that specific bits of information be located and accurately accessed. Accordingly, some of the most important empirical evaluation questions will relate to the performance trade-offs that may exist as a result of this interaction cost. For instance, it is anticipated that, depending on the tasks required of the user, the DB mechanization may pose significant occlusion cost, but information access should be quite quick. Artifacts of this presentation approach may make it difficult to maintain situation awareness of multiple entities compared to the spatially-constrained (SC) variations. The following are some the functions that are supported by the interface (Figure 8):

Direct selection:

The upper and lower bezel of the SC window will be segmented into 10 small "soft keys" which are selectable by the user. Hovering the mouse cursor over the soft keys results in their graphical expansion so that selection is easier than the original size of the keys. Again, this design is conceptually similar to interaction with a multi-function display. Each of the bezel keys correspond with the respective 9-line element (Table 1). The 10th key is reserved for any "remarks" information that the entity 9-line may include. When the subject selects one of the numbered keys on the bezel (via a mouse-driven cursor left-click), the corresponding 9-line text will be presented dynamically via the appropriate access technique being evaluated (access format condition levels of the independent variables).

Run:

This function is activated by the user left-clicking within the area of the SC presentation window. When activated, the 9-line text associated with the entity of interest will cycle within the drawing area via one of the presentation technique formats. Going from the static label state to the run state, the 9-line text will begin with the first category line and proceed in sequence to the end of the 9-line information and repeat the cycle until some other mode is selected by the subject.

Pause/Stop:

When a SC is actively running, a left-click within the drawing area will "freeze" the dynamic presentation of text and display a static string of text that was visible within the dialog box at the instant the left-click selection was made. If the user takes no other action within some to-be-determined period of elapsed time (approximately 5 seconds), the mode will automatically switch to a "stop" state and only the top level entity label text will be presented within the SC presentation window.

Fast-Forward Function:

For this function, the rightmost vertical bezel component is the logical choice as the target soft key (labeled right facing chevron symbol). A fast-forward action during run or direct selection mode will skip the text presentation to the beginning of the next line in the 9-line sequence. Depending on the initial mode prior to this action, the respective presentation will continue from the new start point. To activate a fast forward, the user need only hover the mouse-driven cursor over the fast-forward soft key and left click the mouse input button.

Reversion:

For this function, the left vertical bezel component is the logical choice as the target soft key (labeled left facing chevron symbol). A reversion action during run mode will begin the 9-line sequence again from the beginning of the present text line if selected toward the last 2/3 of that line. If selected during the beginning of the text line (first 1/3), the reversion action will snap to the previous line and begin to run. To activate a reversion, the user need only hover the mouse-driven cursor over the reversion key and left click the mouse input key.

Feedback/completion meter:

A shade fill technique will be used to provide the user simultaneous orientation and progress information. Within each of the numbered bezel keys a left to right fill animation will indicate which line is being presented and, the amount of fill will indicate progress within that line text string. For instance, a half-filled key indicates half of that line of text has been presented and half is yet to be presented. Line presentation completion will correspond closely with the key being completely filled. When presentation switches to the next line, the fill animation will begin again for that line key. This functionality should provide a means for the user to become quickly oriented to location within the 9-line text. Also, this approach should provide a smooth and unobtrusive cycle progress tracking capability.

Presentation Window Entity Dependence:

When the text presentation window of a given entity is selected, the way it (the window) behaves after that may have a significant effect on performance. The proposed research intends to objectively evaluate what, if any, performance effect is dependent on window behavior. This will be treated as a separate independent variable within the study and there will be two levels of the variable. One level acts as if the text presentation window is “attached” to the entity in the way depicted in Figure 8. The location of the text presentation window appears to the upper right of the entity and is tethered to the entity in a way that it will move within the x and y dimensions of the allowable scenario area. For this level the entity dependence variable is called Dynamic Dependence (DD). The second level of entity dependence is fixed to the x and y coordinates of the scenario area once the text presentation window is selected by the user. This is called Static Independence (SI). The way the user interacts with these mechanizations is described the paragraphs below.

Dynamic Dependence (DD) Interaction Mechanization:

The term "dynamic" indicates that the text access window is free to move within the x and y coordinates of the boundaries of the scenario area including exit from the scenario area. The term "dependence" indicates that the text access window moves in conjunction with its parent entity in such a way that it is tethered to the upper right side of the entity as shown in Figure 8. Text access interaction (see the Interaction and Feedback Concepts section earlier in this chapter) is controlled by the subject via a combination of hovering the mouse cursor over the entity data tag and associated multi-function buttons. Activation of the interaction control features is accomplished via a left mouse click. Multiple entity text access data tags can be activated at any particular time by moving the mouse cursor among the entities present at any particular time.

Static Independence (SI) Interaction Mechanization:

The term "static" indicates that the text access window becomes locked to a specific x and y coordinate within the boundaries of the scenario area when activated by the subject. The term "independence" indicates that, once activated, the text access window is independent of the continuous movement of its associated entity. The entity is free to exit the scenario area while any activated text access window remains active.

Activation of the interaction control features is accomplished via a left mouse click. Multiple entity text access data tags can be activated at any particular time by moving the mouse cursor among the entities present at any particular time. A fast double click (within 750 ms) of the left mouse button activation within the text access window acts to "snap" the data tag back to its originally associated entity. This includes removal from the scenario area if the original associated entity has exited the scenario area while the text access window was active.

Occlusion and Layering:

Relative to each other, all the entities, data tags, and text access windows are opaque. To assign occlusion, a layering approach is used based on the order in which entity interaction was sequenced. The most recent interaction brings those objects (set combinations of entities and their associated data tags and activated text access windows) to the top layer and sends each other object one layer back from its previous layering sequence. Similarly, order of entity presence determines layer position in the absence of user interaction. The most recent object with the most recent interaction is at the top or "front" layer. The occlusion sequence described above is the same regardless of whether the text access interaction mechanization is DD or SI.

Experimental Design:

The proposed study is a 4 by 2 full-factorial within-subjects experimental design. There are 4 levels of text access format (DB, RSVP, TS, and CC) and two level of interaction mechanization (DD and SI) independent variables of interest (Table 2). It is intended that order effect confounding for the repeated measures variables will be mitigated by a counterbalancing approach such as a Latin Square design described by Keppel (1982). The General Liner Model Analysis of Variance (ANOVA) F-test seems an appropriate analysis performed separately for the response time and accuracy dependent measures. Post hoc analysis using Tukey honest significant difference (HSD) test will be performed for comparisons where statistically reliable main effects and/or interactions are indicated. A potentially interesting way to analyze effectiveness given the interdependency between response time and response accuracy is to perform an analysis of the independent variables using the response duration measures while holding the accuracy variable constant. The question being, what is the effect on response time when accuracy is always 100% correct?

Table 2. Independent variables matrix.

Text Access Formats				**Interaction Mechanizations**
DB	RSVP	TS	CC	DD
DB	RSVP	TS	CC	SI

<u>Secondary-Task Paradigm for Performance Measurement</u>:

Kantowitz and Sorkin (1983) give a very simple but logical explanation for human information processing mechanism upon which the secondary task performance measurement methodology is based. Central to this concept is that humans have a limited capacity to process information and perform tasks. This is particularly true when tasks require similar modalities or processing channels for their performance (Wickens, 2008). When there is a requirement that two tasks be performed simultaneously, humans can perform the tasks without decrement provided the “pool” of resources available for the effort is not exceeded by the difficulty of the effort (Kahneman, 1973). If the effort to perform two tasks simultaneously exceeds the available resources, some performance decrement across the tasks is expected. If one of the tasks is prioritized by the human as primary, it follows that any performance decrement present will be isolated to the secondary task. As a methodology, this well-established phenomenon can be used to measure any differential capacity requirements of the primary task which may be attributable to an experimental manipulation. For the proposed research, given no other systematic differences other than a manipulation of the text access formats (DB, RSVP, TS, and CC using DD or SI), any significant performance decrement on a secondary task can be interpreted as a result of different levels of “usability” existing among the levels of the text access independent variables.

Kantowitz and Sorkin (1983, p. 184) explain it in very simple terms like this:

> "…we will assume some arbitrary capacity demand for our primary task. The primary task has two levels: 6 bit per second and 8 bits per second. The human channel is assumed to be limited to 10 bits per second. Therefore, either version of the primary task can be performed alone without decrement. We now add a secondary task that requires 4 bit/sec. If we assume that people first allocate their limited capacity to the primary task and then what is left over to the secondary task, we can generate predictions for the limited-capacity model." They continue: "The easy version of the primary task can be successfully combined with the secondary task since both tasks together require 10 bits/sec and this is the limit for our hypothetical human channel. But when the more difficult primary task is combined with the secondary task a total of 12 bits per second is required to perform both tasks. This exceeds available channel capacity so that performance will decline."

The proposed research will use the secondary task methodology to detect any measurable performance differences among the text access formats and entity dependence levels by recording subject secondary task performance that is common across all combinations of the independent variables of interest (Table 2). By instruction and training, subjects will be required to treat the primary task where they will be using the text access formats as their performance priority. They will be instructed to perform the secondary task to the best of their ability without sacrificing performance on the primary task.

The following paragraphs will describe the proposed primary and secondary tasks and their associated performance measures. The tasks were designed to mimic those which are required of some operational command and control tasks (i.e. air traffic control radar, battle management information display, flight deck tactical situation displays).

Primary Task:

The proposed primary task is designed to mimic an operational task where it is necessary for the operator to confirm specific information contained within the 9-line data associated with an identified entity of interest. Metaphorically, the task is representative of a simple query task where a superior communicates an information request to a system operator. In real-world operations, of course a communication like this could take many forms and perhaps is most likely to be auditory/verbal in nature. For the proposed research, it is desirable that the modalities involved in the task (input and output interaction) be as similar as practical and highly controlled. For these reasons, the information query task takes the form of a text-based probe. The query contains the identification of the entity of interest, the 9-line category where the information is contained, and the specific data level information to be confirmed. Figure 9 is an illustration of how a text probe query will be presented to a subject.

When an information query is presented to the subject, the subject's task is to access the appropriate information for the identified entity of interest and confirm that the information associated with the entity is an accurate match (mouse over and select the "Yes (Y)" input) or that the information associated with the entity is not an accurate match (mouse over and select the "No (N)" input). A small input box is presented so the Yes/No input can be logged. Two buttons within the box are labeled with a "Y" for reporting a Yes response and an "N" for reporting a No response. When the cursor is hovered over one of the Y/N buttons, the button nearest the cursor tip is highlighted. Whatever button is highlighted when a mouse button is clicked is recorded and logged as associated input. Again, Figure 9 includes an illustration of confirmation input selection Y/N buttons (query report buttons). The total number of positive and negative but accurate responses will be balanced across the trial blocks while presentation query response type will be randomized.

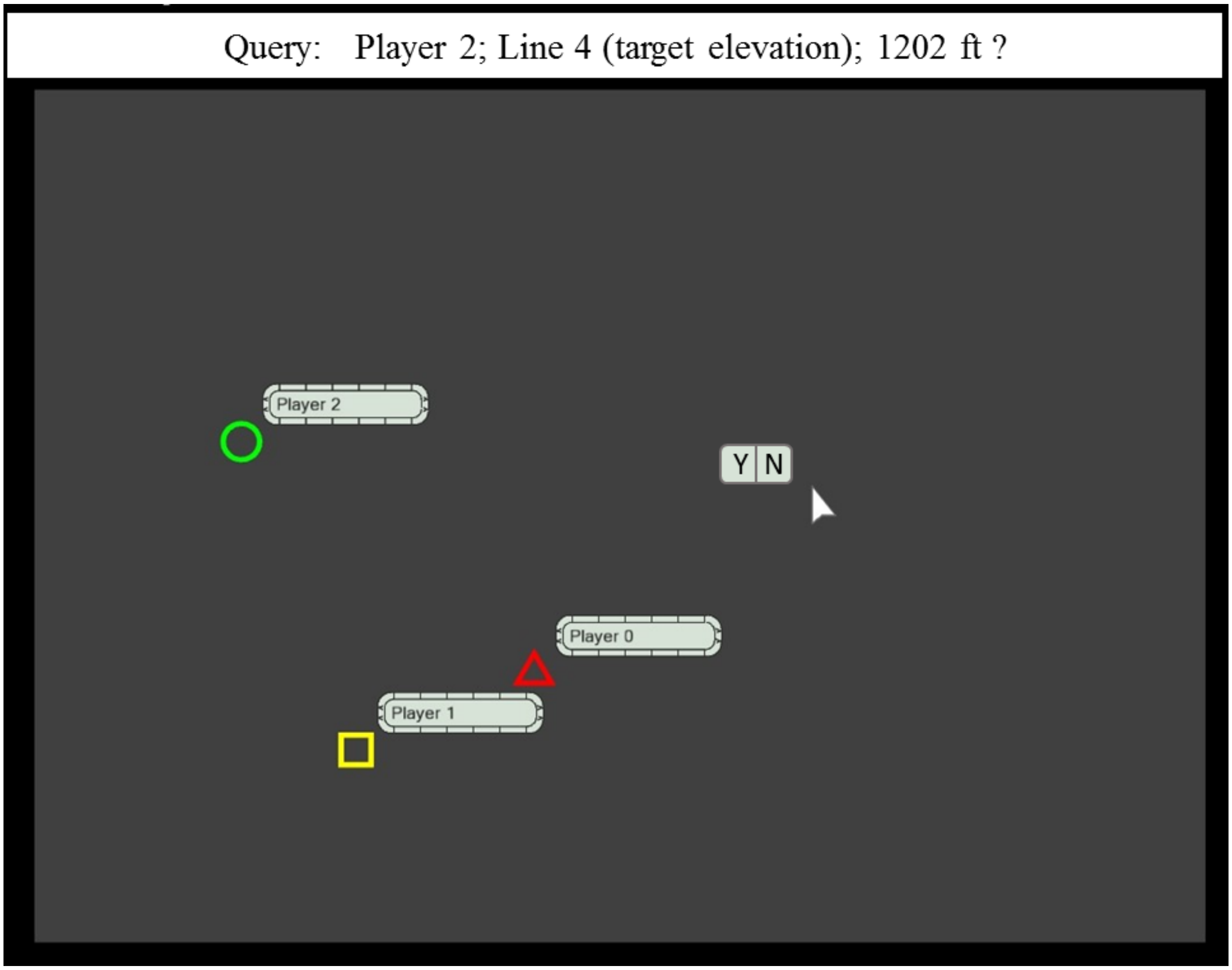


Figure 9. Example primary task query probe.

By explicit instruction, subjects will be requested to treat the primary task as their highest priority task when presented with a query probe. Further, to avoid unwanted speed/accuracy tradeoffs, the instructions will indicate prioritization of accuracy over speed within performance of the query task. Subjects will be directed to perform the secondary task only to the extent possible given spare capacity to do so during accomplishment of the primary task.

The required actions of the subjects to perform the primary and secondary tasks will be consistent and common across the manipulation of the independent variables of interest depicted in Table 2.

Nine (9)-Line Information:

The 9-line data generated for the proposed research was selected so that it meets specific criteria. The data values are simple but realistic and made to appear to be what would be expected of operationally relevant information as if the entities were individual aircraft with tactical assignments. The data are formed of latitude, longitude, and elevation coordinates, heading directions, distances, target types, locations, etc. The data were selected to be easily and quickly understood as well as easily held within the limitations of working memory by minimally trained naïve subjects. No specific skill set is required to perform the tasks. Further, the number of individual text characters within a query probe data value is designed to be accurately perceived by the subject within the limitations of a single foveal fixation. This is based on the normal human ability to perceive 12-14 characters in a single fixation with acceptable accuracy (Robeck & Wallace, 1990; Just & Carpenter, 1998; Rayner, 1998; Öquist, 2006).

The data given in Table 3 is the basis for both the query input and the 9-line content associated with the specific entities to be displayed via the four different formats and two different levels of the dependency variable. For a positive or “Yes” confirmation, the query and entity 9-line data items will match. For a negative or “No” confirmation response, the query item will represent the correct category of information in terms of line number and description but selection logic will include the limitation that the data item origin entity and the query entity cannot be the same thus, the data items are the appropriate type but different values. Subjects will not be able to determine a correct query response by a recognition that the query data item is the wrong category of information for the query line number and description. A correct response includes confirmation of matching data items when the query entity and the selected entity are also the same. Similarly, a correct response is recorded when the subject selects “No” when the query entity are the same but the data items do not match.

Table 3. Nine (9)-line data content.

Line 1	**Line 2**	**Line 3**	**Line 4**	**Line 5**	**Line 6**	**Line 7**	**Line 8**	**Line 9**	**Line 10**
Initial Point	Heading	Distance	Target Elev	Target Desc	Target Loc	Type Mark	Friendly Loc	Egress	Remarks
38.37/81.59	North	163nm	1202ft	Runway	40.49/80.23	Visual	8nm NW	South	KCRW-KPIT
42.40/83.01	Northeast	521nm	626ft	Fuel Tank	42.40/83.01	Laser	5nm NE	Southwest	KDET-KBUF
39.04/84.66	East	413nm	896ft	Hangar	39.04/84.66	Infrared	8nm SW	West	KCVG-KRIC
36.19/95.88	Southeast	230nm	677ft	Landing Pad	36.19/95.88	Beacon	6nm NE	Northwest	KTUL-KLIT
36.08/98.15	South	239nm	2181ft	Bunker	36.08/98.27	Talk on	5nm S	North	KLAS-KNYL
33.63/84.42	Southwest	147nm	1026ft	Tower	33.63/84.42	Radio	7nm S	Northeast	KATL-KMGM
38.03/84.60	West	316nm	979ft	Tank	38.03/84.60	Auto	4nm W	East	KLEX-KSTL
39.90/84.21	Northwest	241nm	1009ft	Launch Pad	39.90/84.21	Code	9nm N	Southeast	KDAY-KORD

Query Event Control:

Query events will occur according to a controlled (manipulated) schedule. Per 10 minute (600 s) data collection session (each a single trial), there will be a total of 22 queries (approximately one every 25 s). The first 25 s period of a trial will not include a query so that enough time is allotted to allow the scenario to be fully populated with secondary task entities prior to the initialization of the first query task. Additionally, a “blank period” of 25 s will occur at the 300 s point of the trial to help avoid predictable query occurrence timing. The initialization of each query will be randomized within the first 5 s of each query period. This also will result in less predictable query timing. The remaining 20 s of each period is allotted for completion of the query task. If the task is not completed within that 20 s window, the query task will timeout and be recoded as such.

Within the 22 queries, half (11) of the correct responses will be data matches (“Yes” response) and half (11) will be data mismatches (“No” response). The selection of the probed entity will be approximately divided equal across the entities within the scenario (22/4 = 5.5 queries per entity per trial). The selection of the query entity will be randomized without replacement and further logical constraint (rule 1) so that the selected entity is present at the time of query and is not scheduled to exit the scenario within the following 10 seconds (rule 2).

Primary Task SysML Activity Diagram:

Figure 10 depicts the activity elements subjects will encounter during performance of each instance of the primary task. The initialization point in the diagram represents that the secondary task was interrupted by a query probe which appeared according the schedule discussed in the previous section. The first step of the activity is for the subject to select the correct entity called for via the query (the selection is performed via a left mouse click after hovering the mouse cursor over the desired entity). Until the correct entity is selected, the activity remains within a loop at this step as depicted in the diagram. Similarly, the following step is for the subject to access the correct line identified as containing the value to be compared to queried value. Depending on the text access technique chosen by the subject, a left click on the proper multifunction soft key will result in a dynamic jump to the associated text line. Once visible, the subject is able to compare the entity text line value with the query value to make a “Yes” input to indicate a match or a “No” input to indicate a mismatch. These selections are also made via a left mouse click. After this input is completed or 20 s has elapsed, the instance of the primary task is terminated.

After termination of the task, the diagram indicates that secondary task performance is to be continued. The secondary task events are persistent during performance of the primary task and the subject has the choice to shift between performance of the primary and secondary tasks at will. By instruction and training, subjects will be encouraged to prioritize performance of the primary task when a query probe is active. This may include abandoning performance of the secondary task while a primary task is underway.

The allowance to shift between the tasks is depicted in Figure 12 which shows the secondary task to be an interruptible activity.

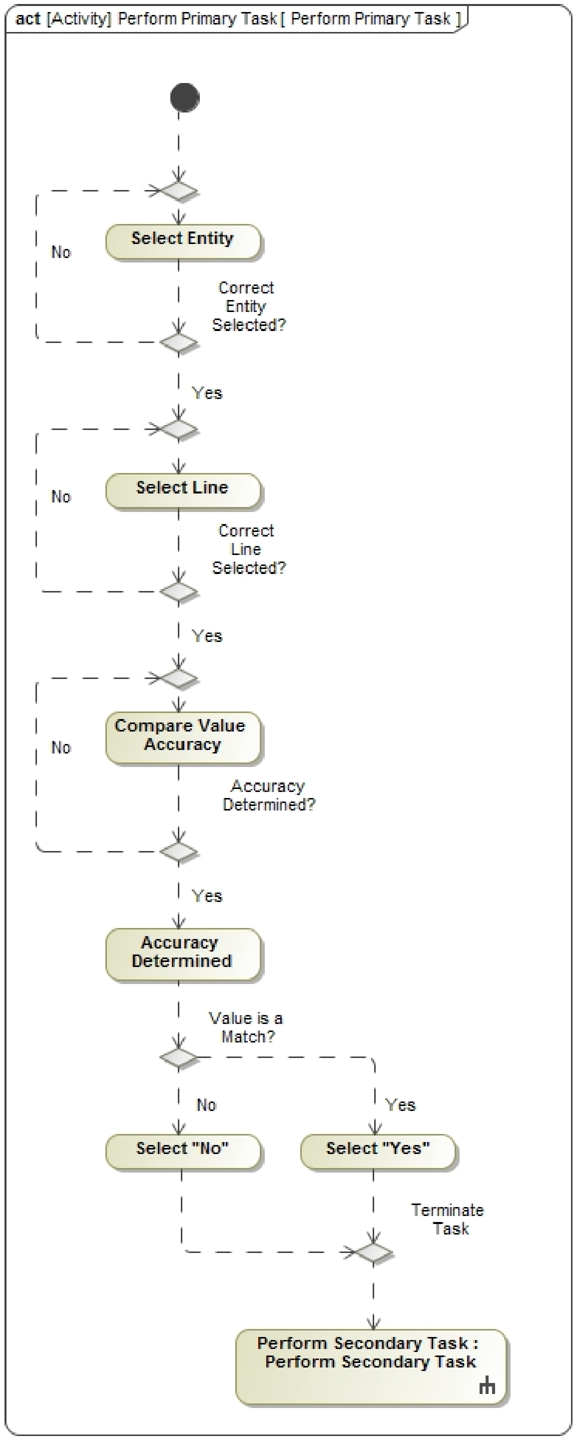


Figure 10. Primary Task SysML Activity Diagram.

Primary Task Dependent Measures:

The primary task dependent measures of interest (within each cell of the balanced test matrix) for this research include response accuracy and response time. The accuracy measure for the primary task is simply a correct response (confirmation that the query and target data values match when they are supposed to and confirmation that query and target do not match when they are not supposed to), or an incorrect response (indicating that the data match when they in fact do not or indicating the data do not match when in fact they do). These responses will be recorded as hits, correct rejections, false selections, and false rejections.

When a query probe is presented, the system clock count will be initiated and the following duration categories will be recorded based on cursor location and mouse click: time to entity text tag selection (correct or otherwise), time of text access interaction initiation, and time to task completion. These measures can be combined to give the elapsed response time of any combination of the intervals.

The following primary task dependent measures will be recorded for purposes of findings analysis and findings interpretation:

1) Accuracy measures: correct, incorrect, and incorrect type.
2) Response time measures: total response time per query task, elapsed time from query presentation to entity selection.
3) Timeout occurrences.
4) Subjective Workload: recoded at the session/trial level (text access format with entity dependence).
5) Text access technique subjective preference recorded after the subject has experienced all of text access formats.

Secondary Task:

The secondary task is designed to mimic a generic scenario monitoring task performed by operators such as air traffic controllers or battle management analysts. The task requires the subject to monitor the activity of several entities and report observed entity status changes when they occur. Figure 11 is a screen capture depicting the presence of three entities in the scenario. The entities in the scenario move across the screen at a constant rate in one of four possible directions relative to screen coordinates: up (+Y), down (-Y), right (+X), or left (-X). Entity identification is differentiated via fairly standard shape and color coded symbols: green circle, red triangle, and yellow square. Each type will include its associated entity ID number within the center area of the entity shape. Attached to each symbol is a data tag (Figure 11) intended to provide specific identification labeling and other status information regarding the associated entity. This data tag is a component where the primary task text access occurs.

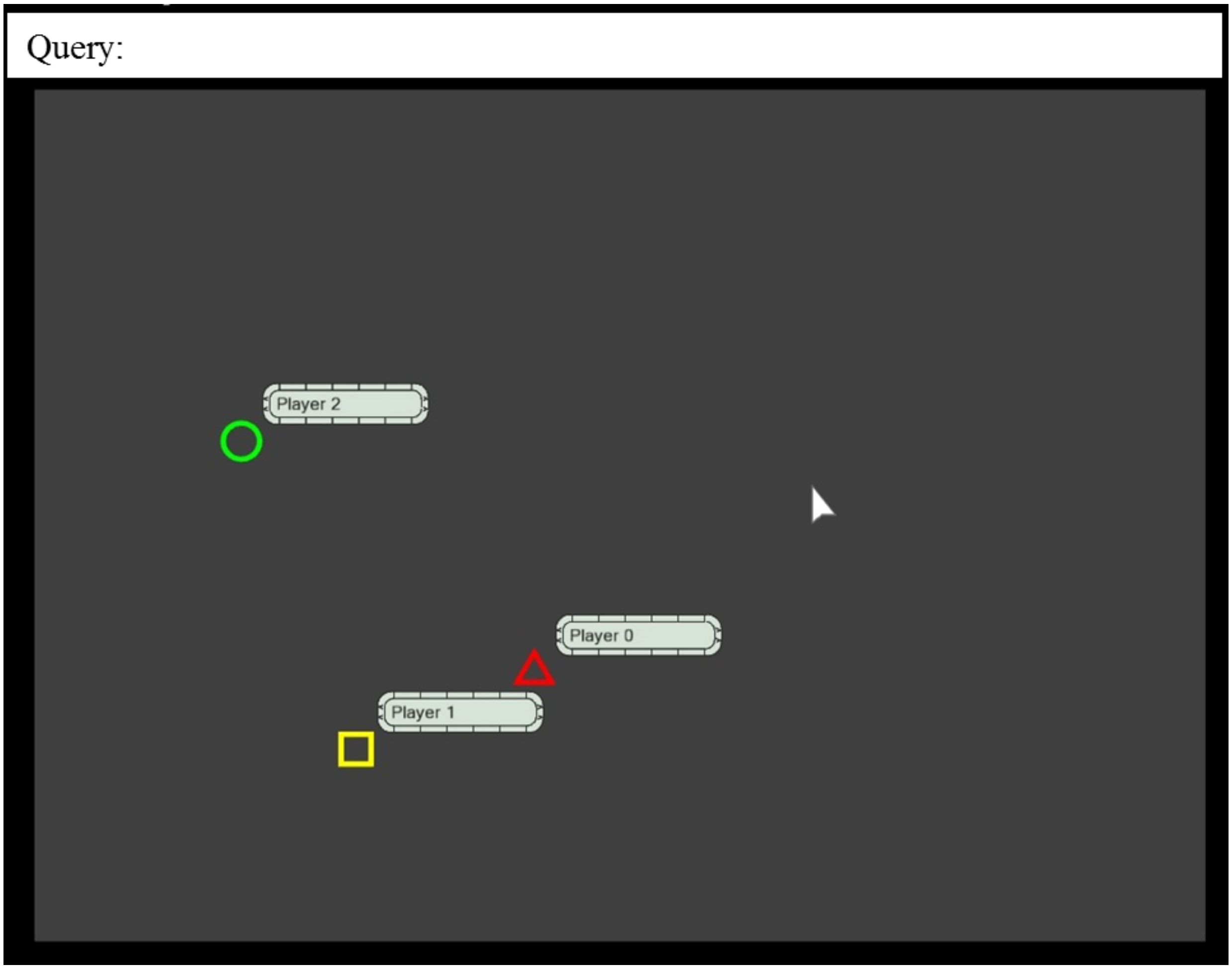


Figure 11. Multiple entity identification symbology, data tags, and cursor.

Subjects will be instructed to monitor multiple entities for three specific events that are "reportable" within the scenario. One reportable event is the entrance of entities into the scenario from "off screen" as they become visible. This event is reported as "presence". Once the entities are present, the task requires that the subject monitor the entities for both changes in direction and changes in identification. At controlled intervals, entities will become present and, once present, the motion of an entity can change from its current trajectory to a +/-90 degree or -180 degree change from the current course. Any time a direction change occurs, it is a reportable event. The final reportable event is a change in entity identification. When this change occurs, the current entity changes shape and color to become one of the other two types of entities. The identification number within the entity shape will not change. Subjects will be instructed to report any time a shape identification change is detected. Table 4 shows the three different reportable entity events which make up the subject activities of the proposed research secondary task. The following paragraphs will describe how the secondary task events are scheduled and how the subjects interact with the secondary task.

Table 4. Reportable entity events for the secondary task.

Entity Change Type		
Presence (P)	Identification (I)	Direction (D)

Secondary Task Event Control:

According to a controlled (manipulated) schedule, players will enter and exit the scenario. No change events occur during the first 15 s of runtime to allow the scenario to be populated with entities before much interaction beyond reporting presence is required of the subject. After that point in time, the scheduling logic loops through all of the active entities that are not currently in an active event status, are not scheduled to exit the scenario within 10 s given their current trajectory, or have had the same event occur within the preceding 10 s. If no entities meet the event criteria, a “blank” event is recorded. For an eligible entity, an event is assigned according to the following probabilities: direction change (p = 0.40), identification change (p = 0.40), and no event assignment (p = 0.20). The frequency of events is controlled by a min/max event timer and thus is controllable with some precision. Similarly, the scenario is designed to maintain a fairly constant mean number of active entities at any given time once the general population is established (at about the first 15 s point). A logical schedule is employed here as well. If the scenario is underpopulated by one, a new entity will be added in 10 s to 20 s. If the scenario is underpopulated by two or more and the next entity is not scheduled to enter for more than 10 s, then an entity is scheduled to enter between 5 s and 10 s. If the scenario logic detects that two entities will exit the scenario with the next 10 s, and no new entities are scheduled to enter within the next 10 s, a new entity will be scheduled to enter the scenario between 2 s and 4 s.

Secondary Task SysML Activity Diagram:

Figure 12 depicts the activity elements subjects will encounter during performance of the secondary task. The initialization point in the diagram represents that the 600 s trial is underway. In general, the secondary task requires the subject to monitor the entities and report on their presence in the scenario as well as changes of identification and direction. The change events occur according to the scheduling logic discussed in the previous section. Subjects select entities and report change events by hovering the mouse cursor over the desired entity and actuating right mouse click to input a response. The diagram indicates that the secondary task activity can be interrupted by a primary task query probe (Figure 10). Subjects will be encouraged by direction and training to prioritize performance of the primary task when the interruptions occur.

Trial termination is indicated when there are no longer any entity events to be reported, there is no active query activity, and the 600 s endpoint has been reached.

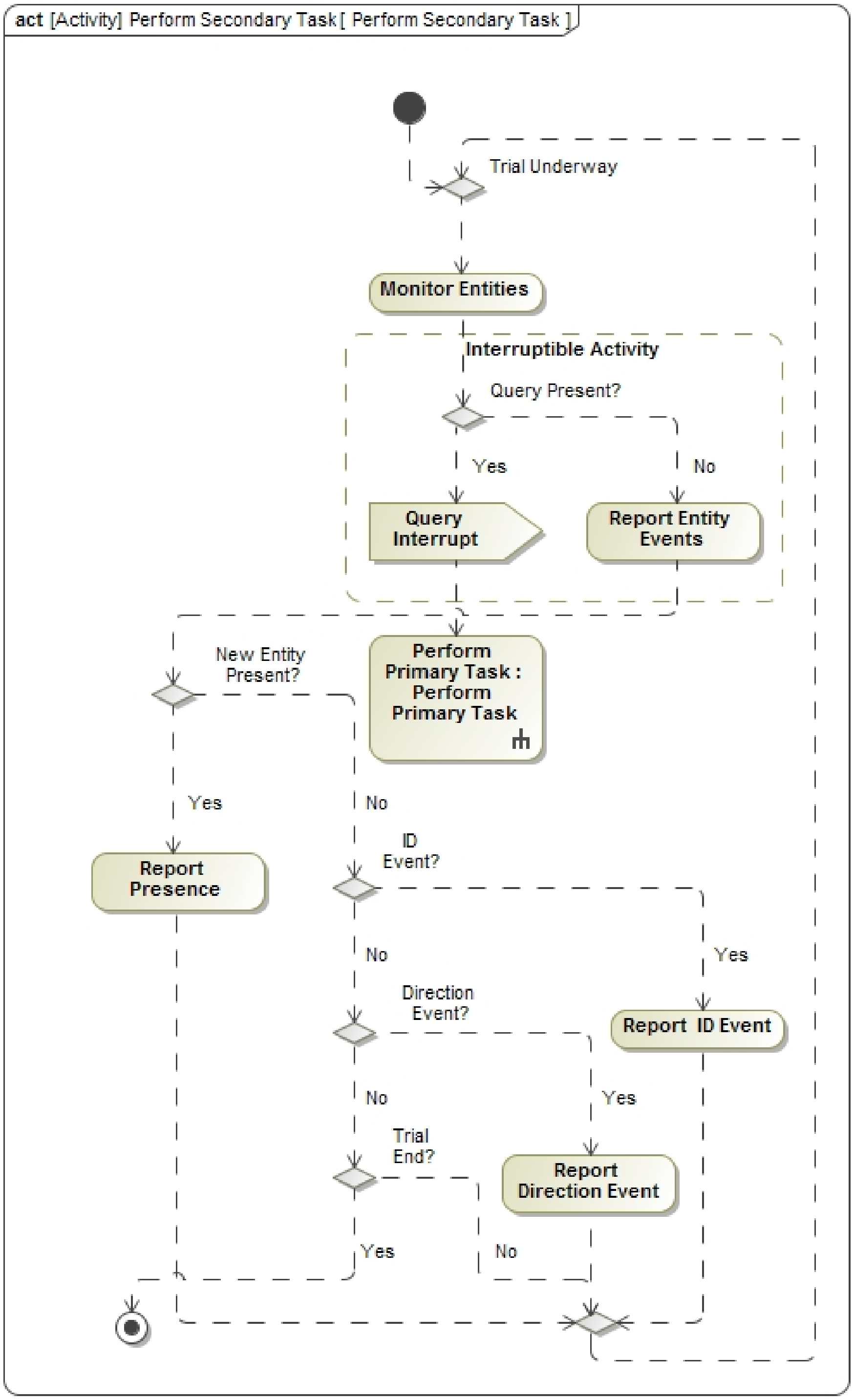


Figure 12. Secondary Task SysML Activity Diagram.

Secondary Task Difficulty:

The intent of the event schedule logic and its associated monitoring functions described above is to generate an appearance of event randomness for the subjects. In fact, the events are highly scheduled and controlled. The desire to control events scheduling is two-fold. First, the basis for a constant set of scenario events can be presented to subjects in a way that they are not able to predict a learned sequence of events. This is accomplished by event initialization dependency on the interaction of the subject and the ability to initialize a common scenario with different initial conditions (first entry direction). Second, difficulty of the secondary task can be manipulated directly by controlling the number of entities (on average) in the scenario and the rate of change assigned to the event schedule probabilities within the bounds of the control criteria described above. Of course difficulty can be manipulated by speeding up or slowing down the entity motion rate as well. It is highly desirable to tune the difficulty of the secondary task so that, when performed in isolation, it is engaging with reasonable effort but near 100% accuracy performance is achievable. This forms the basis of the secondary task measure that is intended to be sensitive to the spare capacity afforded to subjects given the simultaneous performance of the primary task and the manipulation of the independent variable levels within the primary task. This will be discussed in more detail later.

New Entity Initial Condition:

When a new entity is added, the simulation scheduling logic assigns a random direction. The direction is however constrained so that a new entity will not enter from the same side as the previous entity. This is designed to encourage the entities to be spread out. The scheduler then assigns a random position for the entity along the border where it will enter. The monitor does not check to see if any existing entities will overlap. This could potentially make it difficult for a subject to click a new entity but a layering scheme is used so that the new entry is "on top" of the otherwise interfering entity. Similarly, any selected entity is moved to the top layer of the scenario. Due to the speed of the entities and their size, an entity entering from the left or right will not be completely visible before the 10 s event hold expires but presence reporting can be accomplished by the subject's ability to select any visible part of an entity for event reporting purposes. The following section gives more detailed information regarding the actual event reporting process.

Subject and Task Interaction:

All direct interaction between the subject and the scenario is accomplished with a standard "mouse" input device. The mouse is used to move a cursor to any location within the display area within the bounds of the scenario area (defined by the outside border depicted in Figure 11). Proximity of the cursor to an entity automatically associates that entity to the cursor for purposes of selection and further interaction. A "right" click of the mouse device button is used to select the entity for change reporting; an initial left mouse button is used for primary task interaction.

Once an entity is selected via a right click, a small input box is presented at the tip of the cursor arrow so the entity change can be logged. Three buttons within the box are labeled with a “P” for reporting presence of a new entity, an “I” to report an identification change, and a “D” for direction change reporting. Once an entity is selected and the reporting buttons are present, the box remains stationary relative to screen coordinates (fixed in x, y screen coordinates position) until a right or left click is input with the cursor hovered over the P, I, or D button within the “PID box”. When the cursor is hovered over one of the PID buttons, the button nearest the cursor tip is highlighted. Whatever button is highlighted when a mouse button is clicked is recorded and logged as associated input. Figure 13 depicts a static PID box that appeared when “Player 2” was selected via right click. Player 2 has continued to move away from the location where it was selected. The location of the entity selection is approximately where the PID box is drawn. The PID box will remain visible until a P, I, or D button is selected via a mouse click (PID choice condition), a right or left click is made near another entity (re-selection condition), a right or left click is made away from any entity or PID box (cancelation condition), or 10 seconds has elapsed without any click made by the subject (timeout condition). Only one PID box is present at any time during a trial.

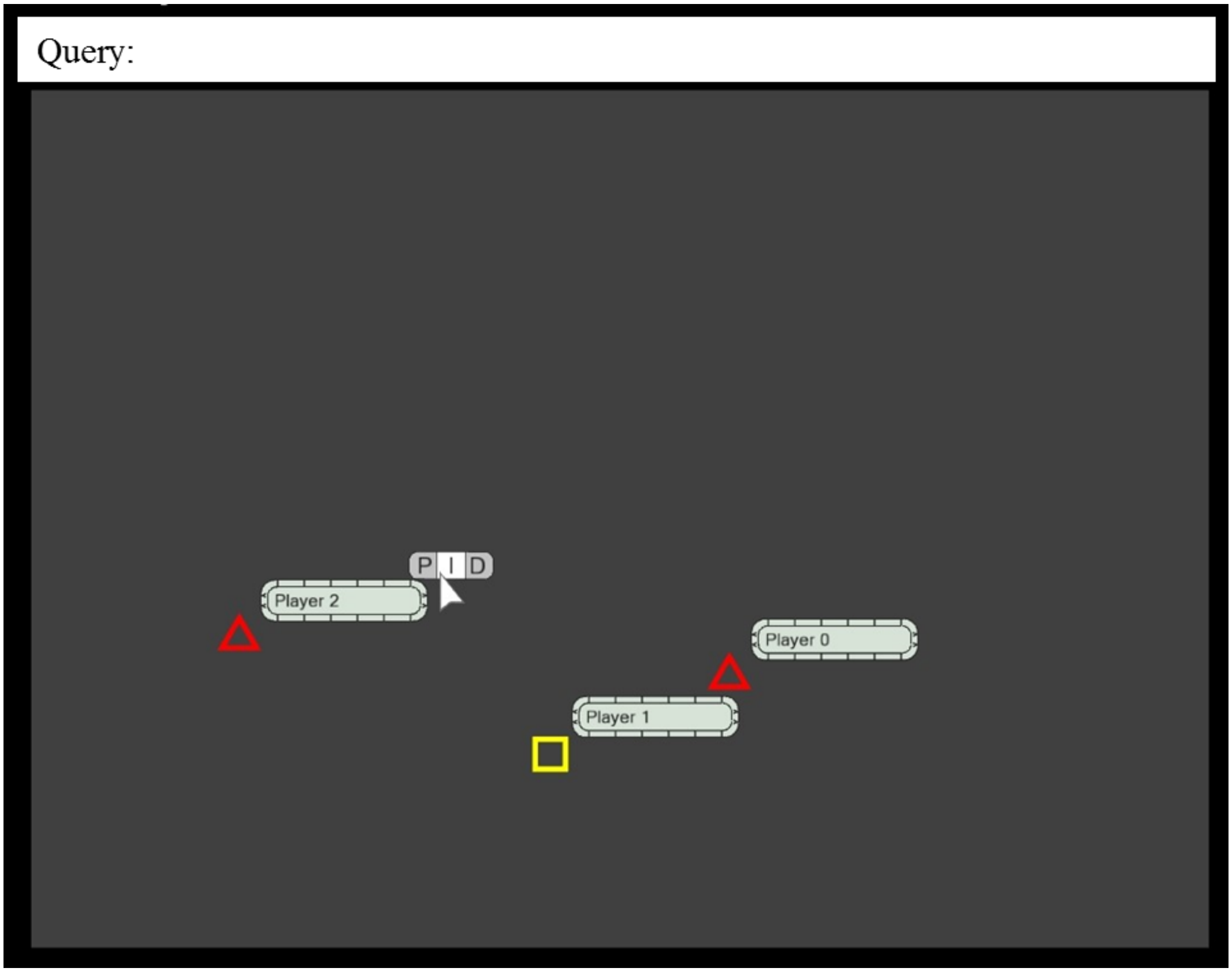


Figure 13. Multiple entities and a “PID box” waiting for participant input.

<u>Secondary Task Dependent Measures</u>:

The secondary task dependent measures are largely focused on accuracy and response time. The response accuracy measures for the secondary task are based on proper selection by the subject using mouse click input and input via the PID box. An accurate input is defined by the selection of the proper entity and PID box button for the preceding change event. Errors are defined several ways: 1) failure to select an entity within the maximum allowable time of a PID event; 2) failure to select the correct PID event change button within maximum allowable time for the correct entity (PID button selection must be made within the maximum allowable time of entity selection or it will be reset); 3) selection of the proper entity but selection of the incorrect PID event button or 4) selection of an entity not associated with a preceding change.

Similar to the primary task response duration measurement, once a change event occurs for a specific entity, a system clock is initiated for that entity. Response time is categorized here as time to correct entity selection (incorrect entity selection is registered as a type of accuracy error), time to a response, correct reply response time, or when timeout is reached. Timeout is the maximum allowable time for registration of a correct response (this also is a type of accuracy error).

The followings secondary task dependent measures will be recorded for purposes of findings analysis and findings interpretation:

1) Accuracy measures: correct entity selection, incorrect entity selection, correct PID selection, incorrect PID selection, and change event misses.

2) Response time measures: total response time per change event, elapsed time from event change to entity selection, elapsed time from entity selection to PID box appearance, and elapsed time from PID box appearance to its removal.

3) Timeout occurrences.

Other Measures:

There are other response measures which may be of interest to ensure a thorough, interpretable, and informative analysis. For instance, cursor activity is an objective measure which is likely to co-vary with the other measures already discussed and it may add to the ability to make inferences regarding the usefulness of the different display formats. Cursor activity as a continuous measure may inform how much work the subject had to perform to achieve a level of accuracy within some response duration. This measure lends itself to investigation in terms of movement area, relative location (to other objects), and rate of input. At the end of active cursor input associated with a change or query event (indicated by its motion and subsequent stop), the location of the cursor on the display screen may be a good approximation of subject instantaneous gaze fixation. This also may be informative in terms of participant task engagement and overall effort.

In general, all events will be recorded with enough framerate resolution to allow the data collection sessions to be replayed (at system clock update rate) along with all system input actions. This will enable the data to be revisited in case new and unanticipated measures are desired for analysis.

Task Training and Feedback:

After completing a written and verbal description of the experimental tasks, Subjects will be trained on the primary, secondary, and combined tasks until the effects of learning on performance become asymptotic. Subjects will be trained first on the secondary task procedures until performance is nearly error free and workload is acceptable. Subjects will be instructed to maintain a continuous visual scan pattern so that change events among the entities are not missed. Next, subjects will be familiarized with the primary task by performing it in isolation of the secondary task. Subjects will experience the primary task with each of the independent variable levels. Once subjects are comfortable with the primary task and performance has stabilized, the primary and secondary tasks will be combined to train simultaneous performance of both the primary and secondary tasks. It will be reiterated that the subjects should prioritize performance on the primary task and, within both tasks, prioritize accuracy performance over speed. Subjects will be reminded to perform the secondary task to the extent possible given any spare capacity to do so.

During the training sessions, subjects will be provided feedback to help inform them firsthand of their accuracy performance on the primary, secondary, and combined tasks. Training performance and feedback will be monitored by the experimenter so that errors can be pointed out and tips for optimal performance of the tasks can be communicated and standardized across subjects. The objective is for all subjects to perform every task as similarly as practical so that their performance is representative of trained operators. This process will also act to ensure that all subjects have the motor dexterity and skill to perform the experimental tasks.

Subjective Feedback Data Collection:

Additional measures of interest are subjective. It will be important to collect subjective workload measures at appropriate points within the overall test matrix. This will give a global measure of participants' perceived level of effort (based on spare capacity) which simultaneously includes primary and secondary task activity. This measure can be correlated with the objective measures. Subjective preference data will also be collected to analyze consistency or lack of consistency with the other measures. It is not uncommon to find user interface comparisons where objective measures are inconclusive or inconsistent with otherwise strong participant indications of preference.

Workload Assessment:

Workload reporting will be performed by having subjects complete a Bedford Workload Scale (Roscoe & Ellis, 1990), decision tree procedure. The Bedford self-assessment workload rating scale is a fairly self-explanatory procedure where subjects report subjective task workload based on consideration of spare capacity to perform additional tasks with regard to the presently performed tasks. Use of the Bedford Scale will be described during subject introduction and instructions. A practice data collection (task training) session will include demonstrating use of the scale. Figure 14 shows the Bedford Workload Scale decision tree.

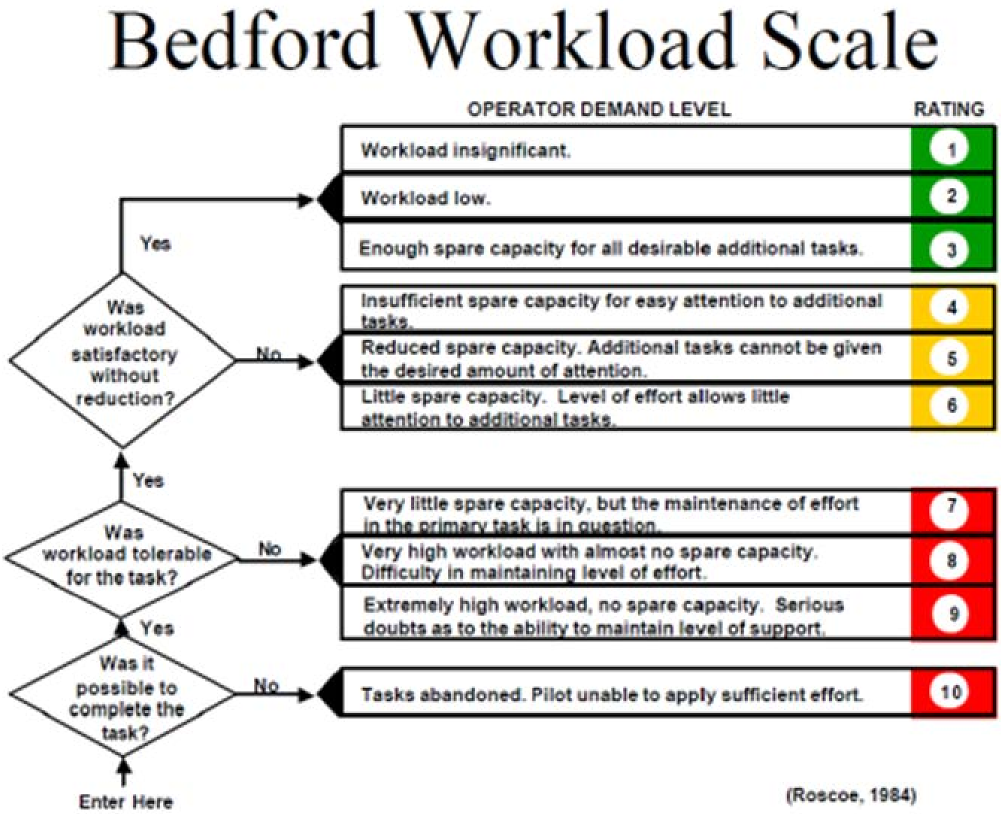


Figure 14. Bedford Workload Scale graphical decision tree.

Subjective Preference Questionnaire:

After completion of all data collection sessions (differentiated by text access format and interaction mechanization), subjects will complete a rating questionnaire designed to record their subjective preferences for the combinations of the independent variables related to the performance of the primary task. Additionally, subjects will be asked to indicate how well they thought they were able to perform the primary task using the combinations of text access formats and interaction mechanizations.

Experimental Control (Minimizing Individual Differences):

A tendency to make speed/accuracy trade-offs is just one of the individual characteristics among the participating subjects that could vary in a way that adds unwanted systematic differences (error) to the response variables within the proposed study. This will detract from the ability to isolate effects in a way that inferences can be attributed to the experimental comparisons of interest (performance differences among the text access techniques and interaction mechanizations). Additionally, there are potential environmental influences on task performance which will be minimized to the extent possible in order to reduce systematic error. This will be accomplished in a way that preserves generalizability of the study findings by allowing the task set to be representative of the defining attributes of real-world applications.

The individual characteristics mentioned above can be thought of as individual differences which participants bring to the data collection session simply because there are differences among humans within any population. Some of these differences lend themselves to experimental control and some do not. Similarly, some are more random than others. Experimental control can take the form of participation selection criteria based on appropriate combinations of physical capability (visual acuity, handedness, etc.), skillset (past experience), training, and/or demonstrated performance (motor dexterity and skill) required for inclusion in the study. When these variables are not controlled directly or where unwanted systematic variably is suspected but is unavoidable, data regarding these differences will be recorded for later analysis. Potential influential conditions which are less controllable exist as well. Examples are participant motivation, pre-existing fatigue, different task completion strategies, and differing priorities. The problematic nature of these influences is that they can vary within individuals across the course of participation in the study.

For the proposed study, the intent is that no unusual or specialized skillset is required for participation. Few physical requirements beyond normal or corrected-to-normal visual acuity and reported normal color vision are necessary. Because the tasks will be performed using a conventional workstation display and control interactions, there is little need to control for handedness but these data will be recorded via a demographics reporting procedure. There is no reason to suspect gender or age (unless extreme) differences but these data will be recorded nonetheless. Similarly, a basic notion of past experience will be collected via demographics reporting in the form of recording experience levels and daily exposure to computer workstation type tasks and video game experience (see Appendix A). In terms of motivation, it is not anticipated that any performance-based reward be used. Although, score comparison with other participants can be used to create a sense of competition which may help maintain motivation. Performance feedback provided during training can help to motivate subjects to continue to work toward performance improvement.

In general, acceptable motivation is assumed due to the volunteer nature of participation. Strategy and performance prioritization variability (response accuracy versus response speed) is mitigated through instruction and training. For example, prior pilot data collection using the proposed secondary task (see the Secondary Task Difficulty Determination Study section below) found that it is important to emphasize an entity-by-entity scan pattern strategy instead of utilizing a centralized fixed gaze approach for the task. Participants will be instructed to prioritize the accuracy of their responses over the speed with which their responses are made. Also, in an attempt to mitigate individual skill differences among participants, as well as ensure that they are adequately familiar with the tasks, training to a baseline level of performance will occur prior to formal data collection participation. Time-to-train will be recorded for later analysis. The utilization of a within-subject experimental design is another individual difference mitigation technique. Counterbalancing will be used to avoid order and learning effects. Fatigue and its influence on attention, motivation, and etc. will be mitigated by including proper rest and recovery periods between data collection sessions. A session duration will be used that does not require unacceptable sustained effort.

Differences within the data collection environment can be controlled where it is anticipated that unwanted systematic effects may result. Data collection will occur in the same physical location using the same lighting conditions and workstation (display and input devices). For the proposed research, the participant posture (seated positon at a workstation display with a keyboard and mouse) will be allowed to vary somewhat to represent how each individual's comfort preference would be established naturally for the use of similar input devices for similar tasks (Figure 1). The workstation display (monitor) location will be fixed on the desktop surface.

Data Collection Procedure and Sequence:

Subjects will be seated at a table equipped with the desktop monitor, keyboard, and mouse. The intent is to mimic the environment depicted in Figure 1. The experiment consists of the following sequence: 1) introduction/consent, 2) demographic questionnaire, 3) vison screening 4) task training, 5) data collection, 6) workload assessment, and 7) preference questionnaire.

1) Introduction/safety briefing/consent: Subjects will read a short description of the study rationale and sequence of events. They will then receive a short safety briefing explaining what to do in the event of an emergency. This briefing will include exit location identification, tornado and shelter-in-place locations, and where to assemble if the building is evacuated. Subjects will read the consent form and sign it to confirm their agreement to participate in the study. Subjects will be informed that their participation in the study is completely voluntary but dependent on their reporting normal or corrected-to-normal vision. Subjects will be informed that they may choose at any time to terminate their participation.

2) Demographic questionnaire: Subjects will complete a questionnaire to report past experience with video displays and tasks which are similar to those used in the study.

3) Vision screening: Subjects will be asked to report normal or corrected-to-normal visual acuity and normal color vision. Corrective lens will be used for all phases of participation in the study if worn normally for similar computer tasks.

4) Task training: subjects will be given detailed instructions regarding the task operation, objectives, priorities, and desired performance. Subjects will also be given practice during fully-dynamic task sessions. The practice sessions will include specific instruction for, and hands-on experience with all interactions they will perform during the actual experimental events. The entire range of input and output variations will be covered during this training period. This training will likely last fewer than 40 minutes. During the training session, subjects will be encouraged to stop and ask questions concerning the tasks. During this time, subjects will practice completing the Bedford Workload Scale self-assessment procedure (Figure 14). Once the experimenter and subject are satisfied that the tasks are fully understood and adequate performance has been demonstrated, the data collection session will be initiated.

5) Data collection: Each subject will complete 32 data collection sessions (4 text access formats by 2 interaction mechanizations by replication (4)) which will each last 10 minutes (320 minutes of actual data collection time). The task will consist of the actions described above as the Primary and Secondary Task sections of this proposal.

6) Workload assessment: Workload reporting will be performed by subjects completing a Bedford Workload Scale (Figure 14) decision tree procedure at the completion of each of the discrete 10 minute sessions.

6) Preference Questionnaire: Subjects will complete the preference questionnaire designed to elicit subjective preference regarding the levels of the independent variables included in the proposed study. Once complete, the subject will have completed their participation in the study.

Safety monitoring:

The studies will not be high-risk and specific safety monitoring will not be required.

Confidentiality protection:

All personal information (data containing personal identifiable information) collected during the study will be stored in a locked cabinet in an office that is secured at all times. All electronic files containing personal information will be password protected and stored only on a secure server. Personal information, hardcopy or electronic, shall be held only while the study remains open per the annual reporting requirements of the Institutional Review Board (IRB). All remaining Informed Consent Document (ICD) materials will be delivered to the IRB coincident with completion of the closure report for the effort and all PII will be destroyed through secure methods (shredding, burning, or deletion from the server) prior to submitting the closure report. The only people who will have access to personal information will be the listed investigators, medical monitor or consultant, AFRL Wright Site Institutional Review Board (IRB), the Air Force Surgeon General's Research Compliance Office, the Director of Defense Research and Engineering Office or any other IRB official involved in the review and approval of this protocol.

10. Risk Analysis

Mild eyestrain in the form of stinging or burning may occur after prolonged viewing of the display hardware. This discomfort should be no worse than watching television in a darkened room for a duration similar to this study. Participants will be informed of possible side effects that may occur during the study and given the option to take a break at any time during test sessions or terminate participation at any point.

11. References

Air Land Sea Application Center. (1997). J-FIRE multiservice procedures for the joint application of firepower. TRADOC-MCCDC-NDC-AFDC Air Land Sea Application (ALSA).

Department of Defense (2005). MIL-STD-2525: Department of Defense Interface Standard: Common Warfighting Symbology. Washington DC, Revised Change 1 edition, July.

Federal Aviation Administration (1987). Advisory Circular 25-11 Transport Category Airplane Electronic Display Systems, AC 25-11, Washington DC.

Just M. A., & Carpenter P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychological Review*, 87(4), 329-354.

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Englewood Cliffs, NJ: Prentice-Hall.

Kantowitz, B. H., & Sorkin, R. D. (1983). Human factors. *New York Etc.: John Wiley*.

Keppel, G. (1992). *Design and Analysis: A Researcher's Handbook.* Prentice-Hall, Englewood Cliffs, NY.

Kruk, R. S., & Muter P. (1984). Reading of continuous text on video screens. *Human Factors*, 26(3), 339-345.

Öquist, G. (2006). *Evaluating readability on mobile devices*. Acta Universitatis Upsaliensis. Studia Lingustica Upsaliensia, Uppsala. ISBN 91-554-6745-8
Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 374-422.

Wickens, C. D. (2008). Multiple resources and mental workload. *Human factors*, *50*(3), 449-455.

Robeck, M. C. & Wallace, R. R. (1990). *The psychology of reading: an interdisciplinary approach, 2nd edition*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Roscoe, A. H. & Ellis G.A. (1990). A subjective Rating Scale of Assessing Pilot Workload in Flight: A Decade of Practical Use, A technical Report, Procurement Executive, Ministry of Defence Farnborough, Hampshire.

12. **Attachments**

a. Informed Consent Document (Phase 1 Study)
b. Informed Consent Document (Phase 2 Study).
c. Curriculum vitae of investigators.
d. Questionnaires (Demographics and Bedford Workload Scale).
e. Recruitment email sample (Phase 1).
f. Recruitment email sample (Phase 2).

INFORMATION PROTECTED BY THE PRIVACY ACT OF 1974

Informed Consent Document
For
Secondary Task Difficulty Parameters Determination (Phase 1 Study)

711 HPW/RHCV, WPAFB, OH

Principal Investigators: DR-III/Eric E. Geiselman, 711 HPW/RHCV

1. **Nature and purpose:** You are participating in the Secondary Task Difficulty Parameters Determination study. Your participation will occur at the Visualization Laboratory, Room 306, Bldg 248, 2255 H Street, WPAFB, OH. This study is being conducted in order to determine the appropriate level of difficulty for a multiple entity secondary task being developed as part of a series of studies planned to investigate the utility of different text presentation techniques. The text presentation studies will use a secondary task paradigm for objective performance measurement. Manipulation of the text presentation techniques will form the basis of a primary task. The present study will be performed to determine the appropriate level of difficulty of a secondary task that is engaging yet, when performed in isolation, affords near errorless performance. The results of this study will define the elements and number of task events required to develop a sensitive and standardized secondary task for future evaluation use. The time requirement for each volunteer subject is anticipated to be a total of one (1) facility visit lasting approximately two (2) hours. Rest breaks are included in the estimated total participation time required for the study. A maximum of 10 subjects will be enrolled in this study. To be eligible for participation you must have 20/20 normal or corrected-to-normal vision using both eyes and normal color vision.

2. **Experimental procedures:**
   a. If you decide to participate, the study is comprised of a sequence of six activities. These include 1) introduction/safety briefing/consent, 2) demographic questionnaire, 3) vision screening, 4) task training, 5) data collection, and 6) workload questionnaires. Reading, understanding, and signing this form will satisfy most of the requirements of the introduction/consent activity. The next activity includes completing a simple questionnaire designed to record your previous experience with visual displays and experience with tasks similar to those used in this study. The vision screening activity includes simply being asked if your vision is normal or corrected to normal acuity (20/20) and normal color vision. Task training includes reading a brief description of the task set followed by an experimenter supervised session of hands-on experience with the task set and the different variations of the display you will experience during the study. The training activity also includes an explanation of a self-assessment workload reporting

questionnaire. Data collection includes performance of the task sets for specific durations without any experimenter assistance.

b. Environmental conditions: You will be seated at a comfortable viewing distance prescribed for the monitor you will use. The room lights will be set for conventional workspace luminance levels. One experimenter will be available and on hand at all times during the entire duration of each session.
c. Criteria for discontinuing the experiment: You must report normal or corrected-to-normal 20/20 vision and no known color vision deficiencies to participate in the study.
d. Alternative: You may choose not to participate in this study; participation is voluntary. You may discontinue your participation at any time simply by informing the experimenter or the medical monitor.

3. **Discomfort and risks:** Mild eyestrain in the form of stinging or burning may occur after prolonged viewing of the display screen. This discomfort should be no worse than watching television (for a duration equal to this study) in a darkened room. You are encouraged to take a break at any time during the test sessions or terminate your participation at any point if these symptoms become problematic for you.

4. **Precautions for female subjects or subjects who are or may become pregnant during the course of this study:** There are no additional precautions for female subjects or those that are pregnant during the course of this study.

5. **Benefits:** You are not expected to benefit directly or be compensated monetarily or otherwise, for your participation in this research study.

6. **Alternatives:** Choosing not to participate is an alternative to volunteering for this study.

7. **Entitlements and Confidentiality:**
   a. Records of your participation in this study may only be disclosed according to federal law, including the Federal Privacy Act, 5 U.S.C. 552a, and its implementing regulations and the Freedom of Information Act, 5 U.S.C. Sec 552, and its implementing regulations when applicable. Personal Identifiable Information to be obtained for this study includes only your name. Your name will only be collected on this informed consent form. Your name will not be linked to your data. You will not be identified in any published reports. All data are combined and analyzed in a group. Your name/signature will be on this consent form and it will be stored in a locked cabinet in an office that is locked when not occupied or held electronically with password protection and stored only on a secure server. It is intended that the only people having access to your information will be the researchers named above, the AFRL Wright Site IRB, the Air Force Surgeon General's Research Compliance office, the Director of Defense Research and Engineering office or any other IRB involved in the review and approval of this protocol. When no longer needed for research purposes your information will be destroyed in a secure manner

(shredding), no later than five years after the study is complete. De-identified raw research data collected during this study may be used in the future for development or testing of computational models; however, this will only involve your de-identified behavioral response data. Complete confidentiality cannot be promised, in particular for military personnel, whose health or fitness for duty information may be required to be reported to appropriate medical or command authorities. If such information is to be reported, you will be informed of what is being reported and the reason for the report.

The decision to participate in this research is completely voluntary on your part. No one has coerced or intimidated you into participating in this program. You are participating because you want to. Eric Geiselman, 711 HPW/RHCV, 937-255-8889 or an Investigator listed above, has adequately answered any and all questions about this study, your participation, and the procedures involved. Eric Geiselman (or Investigator) will be available to answer any questions concerning procedure throughout this study. If significant new findings develop during the course of this research, which may relate to your decision to continue participation, you will be informed. You may withdraw this consent at any time and discontinue further participation in this study without prejudice to your entitlements. The principal investigator of this study may terminate your participation in this study if it is in your best interest to do so. If you have any questions or concerns about your participation in this study or your rights as a research subject, please contact AFRL IRB personnel at (937) 904-8100 or AFRL.IR.ProtocolManagement@us.af.mil.

b. Your participation in this study will not be photographed, filmed or audio/videotaped.

STUDY PARTICIPATION AGREEMENT/CONSENT

Taking part in this research study is completely voluntary. Your signature below shows that:

- ***You agree to be in this study***
- ***The researcher has explained the study to you and you have read and understand the information you have been given***
- ***You were given the opportunity to ask questions about the study and all of your questions have been answered to your satisfaction***
- ***You understand that signing this consent does not take away any of your legal rights***
  You will be given a copy of this signed consent form for your records

Volunteer Signature ____Date ______________

Volunteer Name (printed) ____

Advising Investigator Signature ______________________________Date ______________

Investigator Name (printed) _____________________________________

Witness Signature ____Date _______________

Witness Name (printed) ____

Privacy Act Statement

Authority: We are requesting disclosure of personal information. Researchers are authorized to collect personal information on research subjects under The Privacy Act-5 USC 552a, 10 USC 55, 10 USC 8013, 32 CFR 219, 45 CFR Part 46, and EO 9397, November 1943.
Purpose: It is possible that latent risks or injuries inherent in this experiment will not be discovered until sometime in the future. The purpose of collecting this information is to aid researchers in locating you at a future date if further disclosures are appropriate.
Routine Uses: Information may be furnished to Federal, State and local agencies for any uses published by the Air Force in the Federal Register, 52 FR 16431, to include, furtherance of the research involved with this study and to provide medical care.
Disclosure: Disclosure of the requested information is voluntary. No adverse action whatsoever will be taken against you, and no privilege will be denied you based on the fact you do not disclose this information. However, your participation in this study may be impacted by a refusal to provide this information.

INFORMATION PROTECTED BY THE PRIVACY ACT OF 1974

Informed Consent Document
For
A Human Performance Comparison of Text Presentation Techniques
During Information Extraction Using A Dual-Task Paradigm—Phase 2: Human Performance Comparison Study.

711 HPW/RHCV, WPAFB, OH

Principal Investigators: DR-III/Eric E. Geiselman, 711 HPW/RHCV

7. **Nature and purpose:** You are participating in the Phase 2: Human Performance Comparison Study. Your participation will occur at the Visualization Laboratory, Room 306, Bldg 248, 2255 H Street, WPAFB, OH. This study is being conducted in order to investigate the utility of different text presentation techniques. The text presentation studies will use a secondary task paradigm for objective performance measurement. Manipulation of the text presentation techniques will form the basis of a primary task. The time requirement for each volunteer subject is anticipated to be a total of one or two (1 or 2) facility visits lasting approximately 320 minutes of total participation time. Rest breaks are included in the estimated total participation time required for the study. A maximum of 24 subjects will be enrolled in this study. To be eligible for participation you must have 20/20 normal or corrected-to-normal vision using both eyes and normal color vision.

8. **Experimental procedures:**
   e. If you decide to participate, the study is comprised of a sequence of six activities. These include 1) introduction/safety briefing/consent, 2) demographic questionnaire, 3) vision screening, 4) task training, 5) data collection, and 6) workload questionnaires. Reading, understanding, and signing this form will satisfy most of the requirements of the introduction/consent activity. The next activity includes completing a simple questionnaire designed to record your previous experience with visual displays and experience with tasks similar to those used in this study. The vision screening activity includes simply being asked if your vision is normal or corrected to normal acuity (20/20) and normal color vision. Task training includes reading a brief description of the task set followed by an experimenter supervised session of hands-on experience with the task set and the different variations of the display you will experience during the study. The training activity also includes an explanation of a self-assessment workload reporting questionnaire. Data collection includes performance of the task sets for specific durations without any experimenter assistance.

f. Environmental conditions: You will be seated at a comfortable viewing distance prescribed for the monitor you will use. The room lights will be set for conventional workspace luminance levels. One experimenter will be available and on hand at all times during the entire duration of each session.
g. Criteria for discontinuing the experiment: You must report normal or corrected-to-normal 20/20 vision and no known color vision deficiencies to participate in the study.
h. Alternative: You may choose not to participate in this study; participation is voluntary. You may discontinue your participation at any time simply by informing the experimenter or the medical monitor.

9. **Discomfort and risks:** Mild eyestrain in the form of stinging or burning may occur after prolonged viewing of the display screen. This discomfort should be no worse than watching television (for a duration equal to this study) in a darkened room. You are encouraged to take a break at any time during the test sessions or terminate your participation at any point if these symptoms become problematic for you.

10. **Precautions for female subjects or subjects who are or may become pregnant during the course of this study:** There are no additional precautions for female subjects or those that are pregnant during the course of this study.

11. **Benefits:** You are not expected to benefit directly or be compensated monetarily or otherwise, for your participation in this research study.

12. **Alternatives:** Choosing not to participate is an alternative to volunteering for this study.

7. **Entitlements and Confidentiality:**

c. Records of your participation in this study may only be disclosed according to federal law, including the Federal Privacy Act, 5 U.S.C. 552a, and its implementing regulations and the Freedom of Information Act, 5 U.S.C. Sec 552, and its implementing regulations when applicable. Personal Identifiable Information to be obtained for this study includes only your name. Your name will only be collected on this informed consent form. Your name will not be linked to your data. You will not be identified in any published reports. All data are combined and analyzed in a group. Your name/signature will be on this consent form and it will be stored in a locked cabinet in an office that is locked when not occupied or held electronically with password protection and stored only on a secure server. It is intended that the only people having access to your information will be the researchers named above, the AFRL Wright Site IRB, the Air Force Surgeon General's Research Compliance office, the Director of Defense Research and Engineering office or any other IRB involved in the review and approval of this protocol. When no longer needed for research purposes your information will be destroyed in a secure manner (shredding), no later than five years after the study is complete. De-identified raw research data collected during this study may be used in the future for development or testing of computational models; however, this will only involve your de-identified behavioral response data. Complete confidentiality cannot be promised, in particular for military personnel, whose health or fitness for duty information may be required to be reported to appropriate medical or command authorities. If such information is to be reported, you will be informed of what is being reported and the reason for the report.

The decision to participate in this research is completely voluntary on your part. No one has coerced or intimidated you into participating in this program. You are participating because you want to. Eric Geiselman, 711 HPW/RHCV, 937-255-8889 or an Investigator listed above, has adequately answered any and all questions about this study, your participation, and the procedures involved. Eric Geiselman (or Investigator) will be available to answer any questions concerning procedure throughout this study. If significant new findings develop during the course of this research, which may relate to your decision to continue participation, you will be informed. You may withdraw this consent at any time and discontinue further participation in this study without prejudice to your entitlements. The principal investigator of this study may terminate your participation in this study if it is in your best interest to do so. If you have any questions or concerns about your participation in this study or your rights as a research subject, please contact AFRL IRB personnel at (937) 904-8100 or AFRL.IR.ProtocolManagement@us.af.mil.

d. Your participation in this study will not be photographed, filmed or audio/videotaped.

STUDY PARTICIPATION AGREEMENT/CONSENT

Taking part in this research study is completely voluntary. Your signature below shows that:

- ***You agree to be in this study***
- ***The researcher has explained the study to you and you have read and understand the information you have been given***
- ***You were given the opportunity to ask questions about the study and all of your questions have been answered to your satisfaction***
- ***You understand that signing this consent does not take away any of your legal rights***
  You will be given a copy of this signed consent form for your records

Volunteer Signature ____Date ______________

Volunteer Name (printed) ____

Advising Investigator Signature ________________________________Date ______________

Investigator Name (printed) ______________________________________

Witness Signature ____Date _______________

Witness Name (printed) ____

Privacy Act Statement

Authority: We are requesting disclosure of personal information. Researchers are authorized to collect personal information on research subjects under The Privacy Act-5 USC 552a, 10 USC 55, 10 USC 8013, 32 CFR 219, 45 CFR Part 46, and EO 9397, November 1943.
Purpose: It is possible that latent risks or injuries inherent in this experiment will not be discovered until sometime in the future. The purpose of collecting this information is to aid researchers in locating you at a future date if further disclosures are appropriate.
Routine Uses: Information may be furnished to Federal, State and local agencies for any uses published by the Air Force in the Federal Register, 52 FR 16431, to include, furtherance of the research involved with this study and to provide medical care.
Disclosure: Disclosure of the requested information is voluntary. No adverse action whatsoever will be taken against you, and no privilege will be denied you based on the fact you do not disclose this information. However, your participation in this study may be impacted by a refusal to provide this information.

PROTOCOL CURRICULUM VITAE

1. NAME: Eric E. Geiselman
   GRADE: DR-III

2. CURRENT POSITION TITLE: Senior Engineering Research Psychologist
   LOCATION: 711 HPW/RHCV, Wright-Patterson AFB, Ohio

3. EDUCATION:

M.A.	Experimental Psychology	University of Dayton, 1991
B.S.	Psychology	University of Dayton, 1988

ERIC E. GEISELMAN

711th HPW/RHCV
Bldg 248
2255 H. Street
Wright-Patterson AFB, OH 45433-7022

SUMMARY:

Human Factors Researcher and Program Manager specializing in visualization format design and evaluation for military applications. Primary research areas include human information processing, visual symbology development, imagery integration, and advanced data visualization design. Experience includes applied crew resource management process development, concept development, empirical evaluation, flight test, data analysis, findings reporting, formal presentation, airline operations, safety systems, and teaching. Concentrated experience in the design and evaluation of pilot/vehicle interface concepts for tactical aviation applications. Extensive experience utilizing flight simulation and airborne flight test for the purpose of operational evaluation.

CURRENT POSITIONS:

Engineering Research Psychologist, Air Force Research Laboratory (AFRL), 711th Human Performance Wing, Airman Systems Directorate, Warfighter Interface Division, Battlespace Visualization Branch, Information Visualization Section. Wright-Patterson Air Force Base, Ohio. Duties: design and evaluation of advanced visual display interfaces for performance based human information visualization systems. 07/09 – Present.

Adjunct Faculty, Sinclair Community College, Dayton, OH. Duties: Teaching Aviation Safety and Human Factors topic areas. 01/14 – Present.

Adjunct Faculty, Embry-Riddle Aeronautical University, Cincinnati Center. Duties: Teaching Aviation Safety and Human Factors topic areas. 10/04 – Present.

EDUCATION:

- M.A. Experimental Psychology/Human Factors, University of Dayton, 1991.
- B.S. Psychology, University of Dayton, 1988.

PILOT RATINGS AND FLIGHT EXPERIENCE:

- Commercial Pilot: Single- and Multi-Engine Land Airplane – Instrument.
- CL-65 Turbojet Type Rating.
- Private Pilot: Glider and Single-Engine Seaplane.
- Total time: 7000 flight hours. (as of 11/09).

PREVIOUS EXPERIENCE:

Human Factors Engineer, Aptima Human-Centered Engineering, Inc., Cognitive Systems Engineering Division, Dayton, Ohio. Duties: human-automation interaction and interface design. 3/09 – 7/09.

FAR Part-121 Airline Pilot, Comair Delta Connection, Cincinnati Northern Kentucky International Airport. Duties: First Officer. 3/01 – 3/09.

FAR Part-121 Airline Instructor, Comair Delta Connection, Cincinnati Northern Kentucky International Airport. Duties: Airline Qualification Program Crew Resource Management Instructor. Duties: crewmember instruction, curriculum development, and procedures development. 8/07 – 3/09.

FAR Part-135 Pilot, Miami Valley Aviation, Middletown, Ohio. Duties: DC-3 First Officer. 6/00 – 3/01.

Engineering Research Psychologist, Air Force Research Laboratory (AFRL) Visual Display Systems Branch. Wright-Patterson Air Force Base, Ohio. Duties: design and evaluation of advanced visual display interfaces for head and helmet-mounted displays and night vision systems. 11/98 – 8/01.

Senior Human Factors Engineer, Logicon Technical Services Inc. supporting the Air Force Research Laboratory (AFRL) Synthesized Immersion Research Environment (SIRE) Facility, Wright-Patterson Air Force Base, Ohio. Duties: functionality allocation and evaluation of alternative control technologies, multi-sensory displays, and pilot/system adaptive interface concept development for flight applications and virtual environment interfaces. 4/97 – 11/98.

Engineering Psychologist, Naval Air Warfare Center Aircraft Division, Research & Engineering Group, Crew Systems Integration Division, Human Performance Technology Branch, NAS Patuxent River, Maryland. Program management duties: design and evaluation of visual information symbologies for transparent display application in the flight environment. 9/96 - 4/97.

Senior Human Factors Engineer, Logicon Technical Services Inc. supporting the Armstrong Laboratory (AL) Aerospace Vision Laboratory (AVL), Wright-Patterson Air Force Base, Ohio. Duties: experimental design, data collection, analysis, and results reporting. Primary research areas include human information processing, visual symbology development, and the integration of transparent display imagery. 3/90 - 9/96.

Graduate Assistantship - full scholarship, Armstrong Aerospace Medical Research Laboratory supporting visual display research. 8/89 - 3/90.

Graduate Assistantship - full scholarship, University of Dayton Psychology Department. Work involved experimental design, data collection, and apparatus programming for human automatic information processing research. 8/88 - 8/89.

PROFESSIONAL ACTIVITIES AND ACHIEVEMENTS:

- Numerous invited presentations at various national and international technical conferences.
- Contributing member: Air Force Flight Symbology Design Group.
- Contributing member: Tri-Service Flight Symbology Working Group.
- Guest Lecturer: University of Dayton Engineering Department: (2010 – Present)
- The International Society for Optical Engineering.
- The Human Factors and Ergonomic Society (HFES).
- Southern Ohio Chapter of the HFES: Executive Council Member (2011 - 2013), President 2012.
- Contributor: NATO Advisory Group for Aerospace Research and Development (AGARD) Helmet Mounted Displays and Night Vision Goggles research.
- Contributor: NATO AGARD UPT-7 working group.
- International Symposium on Aviation Psychology: Organizing Committee (2009 - Present).
- Air Force Research Laboratory Human Effectiveness Directorate Civilian of the Quarter: March, 2013.
- Logicon Golden Quill Technical Writing Award Finalist: 1993.

SPECIAL QUALIFICATIONS / TRAINING:

- Statistical Analysis Software: SAS, SPSS-X, SigmaStat/Plot, DeltaGraph Pro, and Excel.
- Rapid prototyping software: VAPS.
- Other flight training and experience: Flight instructor, land and seaplane, glider, and aerobatics.

SECURITY CLEARANCE:

- Top Secret.

RESEARCH AND APPLIED EVALUATION EXPERIENCE AREAS:

- Helmet-mounted display (HMD) symbology design and evaluation.
- Information frame-of-reference for flight applications.
- Transparent display media utility (imagery and imbedded information).
- Aircraft primary flight display design and evaluation.
- Flight deck automation applications.
- Visual information processing and mental representation.
- Automatic information processing.
- Safety training.

SIGNIFICANT RESEARCH CONTRIBUTIONS:

- Provided HMD format design candidates to both X-31 High Alpha Post-Stall Maneuvering and Integrated Helmet Audio-Visual System (IHAVS) flight test programs.
- Designer of the "Theta" transparent ownship attitude display (X-31 and IHAVS flight tests).
- Designer of the "reflected" line-of-sight oriented locator line display for tactical target tracking (IHAVS flight test).
- Designed unique helmet-mounted display functionality and information fusion concepts for an emerging technology demonstration (U.S. Navy).
- Developed and tested symbology color-coding strategy for the Helmet-Mounted Sight Plus program.
- Designer of the Non-distributed Flight Reference Display (NDFR).
- Designer of the Airdrop Guidance Display (ADG) in support of precision non-guided airdrop tasks.

SELECT PUBLICATIONS, ETC:

- Geiselman, E. E., Quill, L. L., Cox, N. J., & Dubois, J. A. (2016). Airdrop Guidance Display Format for a Precision Airdrop Application on an Auxiliary Display Equipped Aircraft. *International Journal of Aviation Psychology*. Vol. 25, Issue 3-4, pp. 141-156, DOI: 10.1080/10508414.2015.1162630.
- **Geiselman, E. E.,** Pinkus, A.R., Haggit J. M., and Task, H .L. (2015). Assessment of Proposed Cab Glass Coating for FAA Control Towers. (Tech. Report AFRL-RH-WP-TR-2015-0074). Warfighter Interface Division, Wright-Patterson Air Force Base, OH.
- Cline, J., Arendt, D., Geiselman, E. E., & Blaha, L. M. (2015). Web-based implementation of the Modified Multi-Attribute Task Battery. Poster presented at the *Fourth Annual Midwestern Cognitive Science Conference*, Dayton, OH.
- Geiselman, E. E., Pinkus, A. R., Garrett, J. S., & Task, H. L. (2014). Evaluation of Proposed Cab Glass Coating for FAA Control Towers. (Report No. AFRL-RH-WP-TR-2014-0123). Wright-Patterson AFB, OH: Air Force Research Laboratory.
- McIntire, J. P., Havig, P. R., & Geiselman, E. E. (2014). Stereoscopic 3D displays and human performance: A comprehensive review, *Displays,* Vol. 35, Issue 1, pp. 18-26, ISSN 0141-9382, DOI:10.1016/j.displa.2013.10.004.

- **Geiselman, E. E.,** Pinkus, A. R., Garrett J. S., and Task, H. L. (2013). Evaluation of Proposed Cab Glass Coating for FAA Control Towers. (Tech. Report AFRL-RH-WP-TR-2013-0123). Warfighter Interface Division, Wright-Patterson Air Force Base, OH.
- Geiselman, E. E., Johnson, C. M., Buck, D. R. & Patrick, T. (2013) Flight deck automation: a call for context-aware logic to improve safety, *Ergonomics in Design: The Quarterly of Human Factors Applications*, Vol. 21, No. 4, pp. 13-18, DOI:10.1177/1064804613489126
- Geiselman, E. E., Johnson, C. M., & Buck, D. R. (2013) Flight deck automation: invaluable collaborator or insidious enabler, *Ergonomics in Design: The Quarterly of Human Factors Applications,* Vol. 21, No. 3, pp. 22-26, DOI:10.1177/1064804613491268
- Task, H. L., Pinkus, A. R., & Geiselman, E. E. (2013). Development of a real-world, sensor-aided target acquisition model based on human visual performance with a Landolt C. In I. Kadar (Ed), *Signal Processing, Sensor Fusion, and Target Recognition XXII, Proceedings of SPIE,* Volume 8745 (pp. 874510-1 – 874510-12). Bellingham, WA: SPIE.
- Blaha, L. M., Geiselman, E. E., & Heft, E. (2013). Toward real-time and predictive behavioral measures of workload capacity. AFRL-AFIT Colloquium on Human Machine Systems, Wright-Patterson AFB, Ohio.
- Geiselman, E. E., & Havig, P. R. (2012). Making the case for off-axis ownship attitude symbology…we may not miss it until it's not there. In P. L. Marasco, P. R. Havig, D. D. Desjardins, & K. R. Sarma (Eds.), *Head- and Helmet-Mounted Displays XVII; and Display Technologies and Applications for Defense, Security, and Avionics VI, Proceedings of SPIE Volume 8383* (pp. 83830H-1 – 83830H-9). doi:10.1117/12.919699
- Havig, P. R., McIntire, J. P., & Geiselman, E. E. (2012). Designing the HMD for perfection: a look at the human eye-brain system. *Proc. SPIE 8383, Head- and Helmet-Mounted Displays XVII; and Display Technologies and Applications for Defense, Security, and Avionics VI,* 838307. DOI:10.1117/12.920008.
- Havig, P. R., McIntire, J. P., Geiselman, E. E., & Mohd-Zaid F. (2012). Why social network analysis is important to Air Force applications. *Proc. SPIE 8389, Ground/Air Multisensor Interoperability, Integration, and Networking for Persistent ISR III,* 83891E. DOI:10.1117/12.920006.
- Ludwig, J., & Geiselman, E. (2012). Intelligent pairing assistant for air operations centers. In *Proceedings of the 2012 ACM international conference on Intelligent User Interfaces (IUI '12)* pp. 241-244. ACM, New York, NY, USA, DOI:10.1145/2166966.2167008
- McIntire, J. P., Havig, P. R., & Geiselman, E. E. (2012). What is 3D good for? A review of human performance on stereoscopic 3D displays. *Proc. SPIE 838. Head- and Helmet-Mounted Displays XVII; and Display Technologies and Applications for Defense, Security, and Avionics VI,* 83830X; DOI:10.1117/12.920017.
- McIntire, J. P., Osesina, O. I., Bartley, C., Tudoreanu, M. E., Havig, P. R., & Geiselman, E. E. (2012). Visualizing weighted networks: a performance comparison of adjacency matrices versus node-link diagrams. *Proc. SPIE 8389, Ground/Air Multisensor Interoperability, Integration, and Networking for Persistent ISR III,* 83891G. DOI:10.1117/12.920012.

- Osesina, O. I., McIntire, J. P., Havig, P. R., Geiselman, E. E., Bartley, C., & Tudoreanu, M. E. (2012). Methods for extracting social network data from chatroom logs. *Proc. SPIE 8389, Ground/Air Multisensor Interoperability, Integration, and Networking for Persistent ISR III*, 83891H. DOI:10.1117/12.920019.
- Geiselman, E. E., & Havig, P. R. (2011). Rise of the HMD: the need to review our human factors guidelines. In P. L. Marasco & P. R. Havig (Eds.), *Head- and Helmet-Mounted Displays XVI: Design and Applications, Proceedings of SPIE,* Volume 8041 (pp. 804102-1 – 804102-11). Bellingham, WA: SPIE.
- Havig, P., McIntire, J. & Geiselman, E. (2011). Virtual reality in a cave: limitations and the need for HMDs? In P. L. Marasco & P. R. Havig (Eds.), *Head- and Helmet-Mounted Displays XVI: Design and Applications, Proceedings of SPIE*, Volume 8041 (pp. 804107-1 – 804107-6). Bellingham, WA: SPIE.
- McIntire, J., Geiselman, E., Heft, E., & Havig, P., (2011). How much camera separation should be used for the capture and presentation of 3D stereoscopic imagery on binocular HMDs? In P. L. Marasco & P. R. Havig (Eds.), *Head- and Helmet-Mounted Displays XVI: Design and Applications, Proceedings of SPIE*, Volume 8041 (pp. 804104-1 – 804104-9). Bellingham, WA: SPIE.
- **Geiselman, E. E.** (2009). Applied threat and error management: toward crew-centered solutions. *Proceedings of the 15th International Symposium on Aviation Psychology*, Dayton, OH.
- **Geiselman E. E.**, (2009). The surprise turn-back: simulation is just not the same. *Flight Training*, Vol. 21, No. 1, pp. 58.
- Jenkins, J. C., Havig, P. R., Heft, E., and **Geiselman, E. E.** (2003). Symbology conceptual research and integration prototyping tool (SCRIPT): development and applications. *Proceedings of the 12th International Symposium on Aviation Psychology*, Dayton, OH.
- Jenkins, J. C., Havig, P. R., and **Geiselman, E. E.** (2002). Development of the non-distributed flight reference helmet-mounted display symbology. *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting*. Human Factors and Ergonomics Society, Baltimore, MD.
- Havig, P. R., Jenkins, J.C., and **Geiselman, E. E.** (2002). A comparison of HMD ownship status symbology and frame of reference orientation during two aircraft control tasks. *Proceedings of the International Society for Optical Engineers (SPIE): Helmet- and Head-Mounted Displays VII.* Rash, C.E., and Reese, C.E. (Eds.), pp 365-376.
- Jenkins, J. C., Reis, G. A., Havig, P. R., and **Geiselman, E. E.** (2002). Evaluation of off-boresight helmet-mounted display symbology formats for interpretation of ownship status information. *Proceedings of the International Society for Optical Engineers (SPIE): Helmet- and Head-Mounted Displays VII.* Rash, C.E., and Reese, C.E. (Eds.), pp 328-340.
- Jenkins, J. C., Thurling, A. J., Havig, P. R., and **Geiselman, E. E.** (2002). Flight test evaluation of the non-distributed flight reference off-boresight helmet-mounted display symbology. *Proceedings of the International Society for Optical Engineers (SPIE): Helmet- and Head-Mounted Displays VII.* Rash, C.E., and Reese, C.E. (Eds.), pp 341-355.

- Jenkins, J. C., Havig, P. R., and **Geiselman, E. E.** (2001). A non-distributed flight reference helmet-mounted display symbology for off-boresight viewing: a dynamic task evaluation. *Proceedings of the International Society for Optical Engineers (SPIE): Helmet- and Head-Mounted Displays VI*. Lewandowski, R., Haworth, L.A., Girolamo, H.J., and Rash, C. E. (Eds.), pp 237-250.
- **Geiselman, E. E.**, Havig, P. R., and Brewer, M. T. (2000). A non distributed flight reference symbology for helmet-mounted display use during off-boresight viewing: development and evaluation. *Helmet- and Head-Mounted Displays V*. Lewandowski, R.J., Haworth, L.A., Stephens, W., and Girolamo, H.J. (Eds.), The International Society for Optical Engineering. Bellingham, WA., pp. 272-283.
- Draper, M. H., **Geiselman, E. E.**, Lu, L. G., Roe, M. M. and Haas, M. W. (2000). Display concepts supporting crew communication of target location in unmanned air vehicles. *Proceedings of the Human Factors and Ergonomics Society 44th Annual Meeting.* Human Factors and Ergonomics Society, Santa Monica, CA., pp. 385-.88.
- Frank, D. L., and **Geiselman, E. E.** (2000). Panoramic Night Vision Goggle Flight Test Results. *Helmet- and Head-Mounted Displays V*. Lewandowski, R.J., Haworth, L.A., Stephens, W., and Girolamo, H.J. (Eds.), The International Society for Optical Engineering. Bellingham, WA., pp. 146-154.
- **Geiselman, E. E.** (1999). Development of a non-distributed flight reference symbology for helmet-mounted display use during off-boresight viewing. *Proceedings for the Fourth Annual Symposium and Exhibition on Situational Awareness in the Tactical Air Environment.* Naval Air Warfare Center - Aircraft Division. Patuxent River, MD., pp. 118-126.
- **Geiselman, E. E.**, and Craig, J. L. (1999). Panoramic night vision goggle update. *Proceedings of the 1999 SAFE Symposium.* SAFE Association, Phoenix, AZ.
- **Geiselman, E. E.**, and Post, D. L. (1999). Helmet-mounted display targeting symbology color coding: An air-to-air scenario evaluation. *Helmet- and Head-Mounted Displays IV*. Lewandowski, R. J., Haworth, L. A., Stephens, W., and Girolamo, H. J. (Eds.), The International Society for Optical Engineering. Bellingham, WA., pp. 66-75.
- Post, D. L., **Geiselman E. E.**, and Goodyear, C. D. (1999). Benefits of Color Coding Weapons Symbology for an Airborne Helmet-Mounted Display. *Human Factors*, Vol. 41, No. 4, pp. 515-523.
- Geiselman, E. E. (1999). Practical considerations for fixed wing helmet-mounted display symbology design. Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting. Human Factors and Ergonomics Society, Santa Monica, CA., pp. 1187-1191.
- Geiselman, E. E., Post, D. L., Brickman, B. J., Rogers-Adams, B., Hettinger, L .J., and Haas, M.W. (1998). Helmet-mounted display targeting symbology color-coding: context vs. population bias. *Head-Mounted Displays III*. Lewandowski, R.J., Haworth, L.A., Stephens, W., and Girolamo, H.J. (Eds.), The International Society for Optical Engineering. Bellingham, WA., pp. 15-24.
- **Geiselman, E. E.**, Brickman, B. J., Hettinger, L. J., Hughes, T., DeVilbiss, C., and Haas, M. W. (1998). Methodology for evaluating off-axis helmet-mounted display ownship information.

Proceedings for the Third Annual Symposium and Exhibition on Situational Awareness in the Tactical Air Environment. Naval Air Warfare Center - Aircraft Division. Patuxent River, MD., pp. 175-182.

- Craig, J. L., and **Geiselman, E. E.** (1998). Further development of the panoramic night vision goggle. *Proceedings of the 1998 SAFE Symposium.* SAFE Association, Phoenix, AZ., pp 26-30.
- Fechtig, S. D., Boucek G. S., and **Geiselman, E .E.** (1998). Preliminary results of the effective information fusion for helmet mounted display technologies program. *Proceedings for the Third Annual Symposium and Exhibition on Situational Awareness in the Tactical Air Environment.* Naval Air Warfare Center - Aircraft Division. Patuxent River, MD., pp. 51-68.
- Hettinger, L. J., Tannen, R. S., **Geiselman, E. E.**, Brickman, B. J., Moroney, B. W., Haas, M. W. (1998) Surgical strike: interface design across task domains. *Proceeding of the Forth Symposium on Human Interaction with Complex Systems.* IEEE Computer Society, Washington DC., pp. 131
- Fechtig, S. D., **Geiselman, E. E.**, Warner, N. W. and Warner, H. D. (1997). Helmet-mounted display concepts, Situational Awareness, and Adaptive Automation. *Proceedings for the Second Annual Symposium and Exhibition on Situational Awareness in the Tactical Air Environment.* Naval Air Warfare Center - Aircraft Division. Patuxent River, MD., pp. 89.
- **Geiselman, E. E.** and Tsou, B. H. (1996). Helmet-display resident target locator line symbology: an evaluation of vector length depiction. *Head-Mounted Displays*. Lewandowski, R. ., Haworth, L. A., Stephens, W., and Girolamo, H. J. (Eds.), The International Society for Optical Engineering. Bellingham, WA., pp. 233-244.
- Meador, D. P., **Geiselman, E. E.**, and Osgood, R. K. (1996). Helmet-display symbology development for the JAST/IHAVS flight demonstration. *Head-Mounted Displays*. Lewandowski, R. J., Haworth, L. A., Stephens, W., and Girolamo, H.J . (Eds.), The International Society for Optical Engineering. Bellingham, WA., pp. 39-49.
- **Geiselman, E. E.** and Osgood, R. K. (1995). Helmet-mounted display attitude symbology: an evaluation of compression ratio. *International Journal of Industrial Ergonomics*, 15: 111-121.
- **Geiselman, E. E.** and Osgood, R. K. (1995). Head vs. aircraft oriented air-to-air target location symbology using a helmet-mounted display. *Helmet- and Head-Mounted Displays and Symbology Design Requirements II*. Lewandowski, R .J., Stephens, W., and Haworth, L. A. (Eds.), The International Society for Optical Engineering. Bellingham, WA., pp. 214-225.
- **Geiselman, E. E.** and Osgood, R. K., (1994). Utility of off-boresight helmet-mounted symbology during a high angle airborne target acquisition task. *Helmet- and Head-Mounted Displays and Symbology Design Requirements*. Lewandowski, R. J., Stephens, W., and Haworth, L. A. (Eds.), The International Society for Optical Engineering. Bellingham, WA., pp. 328-338.
- **Geiselman, E. E.** and Osgood, R. K. (1993). Toward an empirically based helmet-mounted display symbology set. *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting. Human Factors and Ergonomics Society*, Santa Monica, CA., pp. 93-97.
- **Geiselman, E. E.** Osgood, R. K., and Biers, D. W. (1993). A Comparison of Three Aircraft Attitude Display Symbology Structures. (Tech. Report AL/CF-TR-1993-0134). Human Engineering Division, Wright-Patterson Air Force Base, OH.

- **Geiselman, E. E.** and Osgood, R. K. (1992). A comparison of three aircraft attitude display symbology structures during an unusual attitude recovery task. Presented at the 18th Annual Mini-Symposium on Aerospace Science and Technology. Dayton-Cincinnati American Institute of Aeronautics and Astronautics, Dayton, OH, March 26th.
- **Geiselman, E .E.** and Osgood, R.K. (1992). A comparison of three aircraft attitude display symbology structures during an attitude maintenance task. *Proceedings of the Human Factors Society 36th Annual Meeting*. Human Factors Society, Santa Monica, CA., pp. 1450-1454.
- Venturino, M. and **Geiselman, E. E.** (1992). Quantifying the goodness of mental representations of spatial relationships. *Proceedings of the Human Factors Society 36th Annual Meeting*. Human Factors Society, Santa Monica, CA., pp. 1363-1367.
- **Geiselman, E. E.** (1991). A comparison of three aircraft attitude display symbology structures. Unpublished thesis, University of Dayton, Dayton, OH.
- Osgood, R. K., **Geiselman, E.E .**, and Calhoun C. (1991). Attitude maintenance using an off-boresight helmet-mounted display. *AGARD Conference Proceedings No. 1517*: Helmet Mounted Displays and Night Vision Goggles. pp. 14-1 - 14-7.
- Eggemeier, F. T., Granitz, A. B., Rogus, T. E., and **Geiselman, E. E.** (1990). Development of automatic processing with alphanumeric materials. *Proceedings of the Human Factors Society 34th Annual Meeting*. Human Factors Society, Santa Monica, CA., pp. 1357-1361.
- **Geiselman, E. E.**, Wolpert, L., and Venturino, M. (1990). Optical flow sensitivity: effects of field of view and view of optical field. *Proceedings of the 6th Mid Central Ergonomics and Human Factors Mini Symposium*. D. W. Biers (Ed.), Dayton, OH.
- Eggemeier, F. T., Granitz, A. B., Rogus, T. E., and **Geiselman, E. E**. (1989). Automatic Information Processing and High Performance Skills Training: Phase 1. (Tech. Report AFHRL TP-89-70). Dayton, OH: University of Dayton Research Institute, and Logistics and Human Factors Division, Wright-Patterson Air Force Base, OH.
- **Geiselman, E. E.**, Delli, P., and Bajpai, P. K. (1988). Effect of calcium salts on ceramic organic acid composites. *Digest of Papers, Seventh Southern Biomedical Engineering Conference*. D.D. Moyle (Ed.), McGregor and Werner, Washington, DC., pp. 24-27.
- **Geiselman, E. E.**, Morris, L. M., and Bajpai, P. K. (1987). Resorbable ceramic amine and vitamin composites for repairing bone. *Digest of Papers, Sixth Southern Biomedical Engineering Conference*. R. C. Eberhart (Ed.), McGregor and Werner, Washington, DC., pp. 182-185.
- Bajpai, P. K. and **Geiselman, E. E.** (1987). Ceramic amine composites for repairing traumatized hard tissues. *Digest of Papers, Sixth Southern Biomedical Engineering Conference*. R. C. Eberhart (Ed.), McGregor and Werner, Washington, DC., pp. 174-177.

PROTOCOL CURRICULUM VITAE

1. NAME: Michael E. Miller
   GRADE: AD-23

2. CURRENT POSITION TITLE: Associate Professor
   LOCATION: AFIT/ENV, Wright-Patterson AFB, Ohio

3. EDUCATION:

Ph.D.	Industrial and Systems Engineering	Virginia Tech, 1993
M.S.	Industrial and Systems Engineering	Ohio University, 1989
B.S.	Industrial and Systems Engineering	Ohio University, 1987

Short CV Michael E. Miller

EDUCATION
PhD, Industrial and Systems Engineering (Human Factors Specialty), Virginia Tech., 1993
MS, Industrial and Systems Engineering, Ohio University, 1989
BS, Industrial and Systems Engineering, Ohio University, 1987

EMPLOYMENT
Associate Professor, Air Force Institute of Technology, Sept 2015 – Present
Assistant Professor, Air Force Institute of Technology, July 2010 – Sept 2015
Human Factors and Systems Engineer, Eastman Kodak Company, Sept. 1995 – May 2010
Human Factors Engineer, T.A.D. Technical Services on contract to IBM, May 1993 – Sept. 1995
Graduate Assistant, Virginia Tech., Sept. 1990 – Oct. 1992

PROFESSIONAL EXPERIENCE
Dr. Miller has 20 years of experience as a human factors researcher, engineer, or group leader.
Much of Dr. Miller's career has focused on research or development of technology for application into consumer or commercial products. During his career he was responsible for significant design elements of the first generation of IBM's laptop computers, Eastman Kodak Company's consumer digital cameras, and early stage research and development of Organic Light-Emitting Diode displays and lighting systems. He has additionally contributed to research in support of second generation night vision display systems.
During his career, Dr. Miller has concentrated on the development and application of models of human performance to support system design. His current research interests at the Air Force Institute of Technology include measurement of human performance, workload adaptive interfaces, and development and application of human visual performance models to display and lighting design.

SELECT AWARDS AND RECOGNITION
Best Paper-Safety and Human Factors, Industrial and Systems Eng. Research Conference, 2016
Best Paper-Human Factors, Industrial and Systems Eng. Research Conference, 2014
Distinguished Inventor, Eastman Kodak Company, 2005
C.E.K. Mees Award for Technical Achievement, Eastman Kodak Company, 2004

AFFILIATIONS
Senior Member, Society for Information Display
Senior Member, Institute of Industrial Engineers
Member, Human Factors and Ergonomics Society
Member, International Council on Systems Engineering

PROTOCOL CURRICULUM VITAE

1. NAME: Eric L. Heft
   GRADE: DR-II

2. CURRENT POSITION TITLE: Computer Engineer
   LOCATION: AFRL/RHCV Wright-Patterson AFB, Ohio

3. EDUCATION:
   B.S. Computer Engineering Wright State University, 1995

Eric L. Heft

EDUCATION Wright State University Dayton, Ohio
Bachelor of Science in Computer Engineering, Dec 1995
Graduated with a 3.14 cumulative GPA.
Graduated with a 3.71 GPA in CS and CEG.

HONORS Dean's list highest honors and honors.
First place in Wright State University's computer programming contest 1990,1991,1992.
Fourth place in national ACM programming contest 1990.

LANGUAGES and SOFTWARE PACKAGES
Assembly, Access, ADA, C, CGI,CG^2, CSH, Excel,
Fortran, HTML, LabView, Lisp, OneSAF, OpenGL, OpenSceneGraph, Pascal, Perl,
SGI Performer, VBA, Visual Basic, Visual C++, Word

PLATFORMS CPM, DOS, MS Windows, UNIX, Linux, SunOS, SGI

EXPERIENCE

05/03-Present Computer *Engineer - System Administrator*
Battlespace Visualization Branch, Warfighter Interface Division
711 HPW/RHCV, 2255 H Street, WPAFB, Dayton OH

Duties: Write computer programs to support the design and evaluation of aviation displays, display symbology, and human visual performance in an applied research setting. Planning, budgeting, staffing, justification, design, and execution of programming projects that support human factors research dealing with cognitive and perceptual aspects of visual information displays. Writing computer programs that support the photometric and colorimetric measurement, characterization, and modeling of electronic visual displays. Use knowledge of the design and applications of aviation displays, visual psychophysics and perception, CIE photometry and colorimetry, and human factors to design and write programs efficiently in an advanced aviation display research environment.

03/97-12/03 (part time)

Web Master / Production Engineer
Lee Lectronics Limited.
785 Alpha Rd, Alpha Ohio.

Editor of the F1A function generator user manual.
Assembled F1A function generators.
Created quality control checklists.
Placed all manuals and ordering information on the web.
Responsible for maintaining LLL's web pages.

03/96-05/03

Lead Programmer/Analyst - System Administrator
Northrop Grumman Information Technology,
AFRL/HECV, 2255 H Street, Wright Patterson AFB , Dayton OH

System Administrator for Silicon Graphics Workstations.

System Administrator for 12 node PC windows and Linux LAN.
Negotiated SGI support contracts.
Proposed, received funding for, developed, and demonstrated portable color control system for the color display laboratory.
Proposed and developed Linux Performer based F15 simulation
Integrated F15 simulator to work with head tracking and new helmet mounted displays.
Designed, developed, and maintained a MS Windows based, OpenGL, aircraft off bore sight symbology demonstrator.
Used F15 flight simulator used to study the effects of symbology and frame of reference for pilots placed in unusual attitudes.
Developed a web based search engine under Perl to retrieve articles from the Performer email archives.
Created many data collection programs to study various effects of motion on target recognition.
Developed a data collection program for the study targets

embedded in dynamic noise.
Generated program suite to perform image processing of over 2700 target images.

06/93-03/96 C/C++ Programmer -
Northrop Grumman Information Technology

2255 H Street, Wright Patterson AFB , Dayton OH

Created a software program that loaded raw sonar data for viewing on the head tracked helmet mounted display.
Diagnosed corruption of 50+ megabytes of binary data transferred from tape and built a program to dynamically reconstruct the original data stream.
Used the Visual Research Graphics development environment to create a suite of programs that display three different types of dynamic noise.
Created a GUI image-processing program to experiment with advanced image processing techniques.
Tested the KHORUS development environment.

11/89-05/93 Teaching Assistant –
Wright State University
3640 Col Glen Highway, Dayton Ohio
Prepare lab project assignments and solutions.
Instruct students on lab assignments.
Responsible for keeping lab equipment operational.

05/87-07/88 Draftsman - Aleck Industries
Responsible for turning Design Engineers rough sketches into finished blueprints.

PROTOCOL CURRICULUM VITAE

1. NAME: Lauren E. Menke
GRADE: Contractor

2. CURRENT POSITION TITLE: Research Coordinator
LOCATION: 711 HPW/RHCV, Wright-Patterson AFB, Ohio

3. EDUCATION:
M.S. Business & Organizational Management Counseling, Wright State University 2012
B.S. Psychology, Wright State University 2009
A.S. Baking & Pastry Arts, Sullivan University 2004

Lauren E. Menke

Education:
M.S. in Business & Organizational Management Counseling, 2012
Wright State University

B.S. in Psychology, 2009
Wright State University

A.S. in Baking & Pastry Arts, 2004
Sullivan University

Professional Training/Coursework:
- SharePoint Owner Training (Completed 2018)
- Specialized Training for Advanced Brain Monitoring X-10 B-Alert; EEG Device (Certified 2012)
- ConforMat; Pressure Tracking Device (Manufacturer Representative Briefing 2014)

Relevant Experience:
Ball Aerospace, 2018 – Present
Program Manager I, (Research Coordinator)
Ms. Lauren Menke currently supports RHCV at Wright Patterson Air Force Base where she coordinates in support for the lab. Her activities include:
HIRT Contract, 2018–present:
- Recruit, schedule, and coordinate research participants
- Run research participants through experimental protocols
- Record Management to meet research project and branch expectations
- Contribute to scientific reports, papers, and presentations

Ball Aerospace, 2016 – 2018
Program Manager I, (Research Exercise Coordinator)
Ms. Lauren Menke supported RHA at Wright Patterson Air Force Base where she coordinates warfighter exercises in support for the lab. Her activities include:
ASIST Contract, 2016–2018:
- Build and maintain detailed exercise schedules
- Run, administer and organize all project activities in cooperation with and under the direction of the Project Manager, aiming at the flawless execution of the project
- Liaise with customer to identify and define project requirements, scope and objectives

Ball Aerospace, 2011 – 2016
Human Factors Research Professional Associate
WIRTO Contract, 2011–2014; HIRT Contract, 2014–present:
- Expertise conducting and training subordinates in research protocol and collecting research data. Expertise includes set-up, calibration, troubleshooting, and operations of electroencephalograph, electrocardiogram, electrooculograph, trans cranial dopler, ConforMat and five different eye tracking data collection equipment
- Supervise quality assurance of data collection and physiological device cleaning and sanitizing.
- International collaboration and participation in research studies abroad including Defence Science & Technology Organisation, Royal Australian Air Force and Defence Technology Agency
- Composed 11 additional items for the Device Comfort Questionnaire to distribute to operators in Exercise Black Skies 2016
- Compile and calculate experimental data for analyses
- Schedule, screen and coordinate participants for multiple studies
- Participate in task and stimuli development for research studies
- Create training slides and checklists for research studies
- Derive systems and checklists for studies

Dr. Jean Edwards' Laboratory, Wright State University August 2008-June 2009
Research Assistant
- Conducted group sessions for participants, explained the experiment, collected consent, accountable for sensitive data
- Collected data and gave credit to participants who completed the survey
- Exhibited strong communication skills
- Acted as a liaison between the head researcher and fellow research assistants and participants
- Performed data collection from O*Net
- Coded and cleaned data for a meta-analysis
- Input data for SAS

Dr. John Flach's Laboratory, Wright State University August 2009-November 2009

Research Assistant

- Worked with a Brain Computer Interface device to investigate the extent to which the individual channels of this device could be brought under conscious deliberate control by the user
- Conducted literature reviews on Brain Computer Interface systems

Professional Societies: Toastmasters International

Publications:

Vieane, A., Funke, G., Mancuso, V., Greenlee, E., Dye, G., Borghetti, B., Miller, B., Menke, L., Brown, R. (2016). Coordinated displays to assist cyber defenders. Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Tolston, M., Strang, A.J., Funke, G.J., Miller, B., Brown, R., Menke, L. (2016). Evaluating the relationship between team performance and joint attention with longitudinal multivariate mixed models. Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Funke, G., Greenlee, E., Carter, M., Dukes, A., Brown, R., Menke, L. (2016). Which eye tracker is right for your research? Performance evaluation of several cost variant eye trackers. Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Tolston, M.T., Finomore, V., Funke, G.J., Mancuso, V., Brown, R., Menke, L., & Riley, M.A. (2016). Effects of biasing information on the conceptual structure of team communications. Proceedings of the Applied Human Factors and Ergonomics International Conference.

Funke, G., Dye, G., Borghetti, B., Mancuso, V., Greenlee, E., Miller, B., Menke, L., Brown, R., & Vieane, A. (2016). Development and validation of the air force cyber intruder alert testbed (CIAT). Proceedings of the Applied Human Factors and Ergonomics International Conference.

Menke, L.E., Best, C., Funke, G.J., & Strang, A.J. (2015). Warfighter acceptance of future physiological monitoring and augmentation: A coalition study. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting.*

Strang, A., Funke, G., Satterfield, K., Miller, B., Menke, L., & Brown, R. (2015). Effects of task-load transitions on EEG coupling in a high-tempo cooperative task: Verifying a basic utility for future team monitoring. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting.*

Menke, L.E., Best, C., Funke, G.J., & Strang, A.J. (2015). A coalition study of warfighter acceptance of wearable physiological sensors. *Proceedings of the International Symposium on Aviation Psychology, 18*, 440-445.

Funke, G., Knott, B., Strang, A., Dukes, A., Miller, B., Brown, R., & Menke, L. (2015, July). *Effects of access to a physiologically-based operator workload monitor on dynamic team workload balancing*. Poster presented at the 6th International Conference on Applied Human Factors and Ergonomics and the Affiliated Conferences, Los Vegas, NV.

Mancuso, V.F., Greenlee, E.T., Funke, G., Dukes, A., Menke, L., Brown, R., & Miller, B. (2015). Augmenting cyber defender performance and workload through sonified displays. *Proceedings of the 6th International Conference on Applied Human Factors and Ergonomics and the Affiliated Conferences, 6*, 2348-2355.

Mancuso, V.F., Funke, G.J., Greenlee, E., Strang, A., Menke, L., Brown, R., Dukes, A., & Miller, B. (2015, April). Off-body operator state detection: Utility of a pressure sensitive seat in detecting workload. Poster presented at the *Health and Human Performance Research Summit*, Dayton, OH.

Russell S.M., Funke G.J., Flach, J. M., Watamaniuk, S. N. J., Strang, A. J., Miller, B. T.,

Dukes, A., Menke, L. E., and Brown, R. D., (2014). Alternative indices of performance: An exploration of eye gaze metrics in a visual puzzle task. AFRL-RH-WP-TR-2014 -0095, Air Force Research Laboratory, Human Effectiveness Directorate, Wright-Patterson AFB, OH

Funke, G.J., Knott, B.A., Mancuso, V., Strang, A.J., Estepp, J., Brown, R., Menke, L., Miller, B., & Dukes, A. (2013). Evaluation of subjective and EEG-based measures of mental workload. In C. Stephanidis (Ed.). Human-Computer Interaction International 2013 – Posters' Extended Abstracts (pp. 412-416). Spring-Verlag Berlin Heidelberg. DOI: 10.1007/978-3-642-39473-7_82

Alarcon, G.M., Edwards, J.M., & Menke, L.E. (2011). Student burnout and engagement: A test of the conservation of resources theory. *Journal of Psychology: Interdisciplinary and Applied*.

A Human Performance Comparison of Text Presentation Techniques

During Information Extraction Using A Dual-Task Paradigm

FWR20170050H 3.01

DEMOGRAPHIC QUESTIONNAIRE

HUMAN PERFORMACNE STUDY
GEISELMAN, MILLER, HEFT, & MENKE
AFRL/711 HPW/RHCV and AFIT/ENV

Subject number ____ Age ________

(circle one):	20/20	Vision Corrected to 20/20	Less than 20/20
(circle one):	Color Vison	Normal	Deficient
(circle one):	Glasses	Contacts	Neither
(circle one):	Left-handed	Both	Right-handed

Hours per week playing video games: _____

If applicable, what type of video games do you play? _____________________________

If applicable, what type of video game system do you use? (circle all which apply):

Computer TV Portable

Hours per week watching TV: _____

Hours per week on a computer: _____

Is computer use for (circle all which apply):

Work Leisure Gaming

Comments:

Recruitment E-mail "Invitation to Participate" (Phase 1)

To: RHCV Personnel
Subject: Study Participation Opportunity

Protocol: FWR20170050H v1.00 approval date range dd mmm yyyy to dd mmm yyyy

Greetings,

We are looking for volunteers to participate in the Secondary Task Difficulty Parameters Determination study. Your participation will occur at the Visualization Laboratory, Room 306, Bldg 248, 2255 H Street, WPAFB, OH. This study is being conducted in order to determine the appropriate level of difficulty for a multiple entity secondary task being developed as part of series of studies planned to investigate the utility of different text presentation techniques. The text presentation studies will use a secondary task paradigm for objective performance measurement. Manipulation of the text presentation techniques will form the basis of a primary task. The present study will be performed to determine the appropriate level of difficulty of a secondary task that is engaging yet, when performed in isolation, affords near errorless performance. The results of this study will define the elements and number of task events required to develop a sensitive and standardized secondary task for future evaluation use. The time requirement for each volunteer subject is anticipated to be a total of one (1) facility visit lasting approximately two (2) hours. Rest breaks are included in the estimated total participation time required for the study. A maximum of 10 subjects will be enrolled in this study. To be eligible for participation you must have 20/20 normal or corrected-to-normal vision using both eyes.

Please let me know if you are interested and I can send you more information. Thank you for your time.

V/R,

Eric Geiselman

____
Eric E. Geiselman
Engineering Research Psychologist
Battlespace Visualization Branch
Warfighter Interface Division, 711 HPW/RHCV
2255 H Street
Wright-Patterson AFB, OH 45433-7022

eric.geiselman@us.af.mil
DSN: 785-8889
(937) 255-8889

A Human Performance Comparison of Text Presentation Techniques

During Information Extraction Using A Dual-Task Paradigm

FWR20170050H 3.01

Recruitment E-mail "Invitation to Participate" (Phase 2)

To: RHCV Personnel
Subject: Study Participation Opportunity

Protocol: FWR20170050H v3.00 approval date range dd mmm yyyy to dd mmm yyyy

Greetings,

> We are looking for volunteers to participate a Human Performance Comparison Study. Your participation will occur at the Visualization Laboratory, Room 306, Bldg 248, 2255 H Street, WPAFB, OH. This study is being conducted in to investigate the utility of different text presentation techniques. The text presentation study will use a secondary task paradigm for objective performance measurement. Manipulation of the text presentation techniques will form the basis of a primary task. The time requirement for each volunteer subject is anticipated to be a total of one or two (1 or 2) facility visits lasting approximately 320 minutes of total participation time. Rest breaks are included in the estimated total participation time required for the study. A maximum of 24 subjects will be enrolled in this study. To be eligible for participation you must have 20/20 normal or corrected-to-normal vision using both eyes and normal color vision.

Please let me know if you are interested and I can send you more information. Thank you for your time.

V/R,

Eric Geiselman

____

Eric E. Geiselman
Engineering Research Psychologist
Battlespace Visualization Branch
Warfighter Interface Division, 711 HPW/RHCV
2255 H Street
Wright-Patterson AFB, OH 45433-7022
eric.geiselman@us.af.mil
DSN: 785-8889
(937) 255-888

Appendix E: IMAGE 2019 Conference Paper

THE DEVELOPMENT AND EVALUATION OF A SPATIALLY-CONSTRAINED TEXT ACCESS METHOD FOR INFORMATION VISUALIZATION

Eric E. Geiselman
Air Force Research Laboratory
WPAFB, OH, USA

Michael E. Miller
Air Force Institute of Technology
WPAFB, OH, USA

ABSTRACT

To maximize visualization effectiveness, it is not uncommon for graphical data to be augmented with alphanumeric (i.e., text) symbols to provide detailed information and define specific values. However, a human performance cost can be associated with current portrayal and access techniques as the resulting portrayal can occlude critical information and significantly degrade operator performance. The current research seeks to develop and demonstrate a method which reduces occlusion and that may result in improved human performance. Quantitative performance measures include both accuracy and response time. The basis for development of the spatially-constrained text access technique is discussed as well as the implementation affordances and limitations. Further, the paper discusses the operational applications and implications of this research.

INTRODUCTION

The fundamental objective of information visualization is to enable human operators to interpret complex relationships among data. The overall goal is to afford accuracy and timely understanding for meaning making and decision support. Information accessibility and interactivity are also desirable across a wide variety of portrayal approaches. These characteristics in proper combination form the overall effectiveness of a visualization design and are strongly related to human operator performance. To realize effective visualization, it is not uncommon for graphical data visualizations to be augmented with alphanumeric (text) symbols for the purpose of labeling, detail presentation, defining specific values, etc.

A good example visualization is that used in Command, Control, Communication, and Intelligence (C^3I) applications. Examples of these types of displays are: air traffic control large-area radar depictions, Airborne Warning and Control System (AWACS) battlespace visualization workstation displays, cyber activity monitoring graphs, and large-area situation displays. Figure 1 shows the type of operator workstation under consideration. Figure 2 is an example of the complexity often found in existing visualizations. Within the Figure 3 example, a “pop-up dialog box” is utilized to provide text-based information regarding entities of interest. In the example, the pop-up dialog occludes a portion of the displayed information.

Figure 1. Command and control operator workstation example.

Visualization requirements for supervisory control of multiple remotely piloted aircraft, space situation awareness, and flightdeck tactical situation awareness all pose similar challenges. For the user, as a decision maker, the needs are common among these examples. The display must support the recognition of the dynamic spatio-temporal relationships within the wide area “big picture”, traditionally supported by a “God’s eye view” of the scenario. The display should also support the operator’s ability to “drill down” to obtain detailed information pertaining to entities of interest. The continuum between these extremes must be accessible as well. Simultaneously, presentation formats must adhere to the principles of readability that have evolved over time. To be able to determine if novel formats work well, empirical validation methods are needed during usability evaluations.[1]

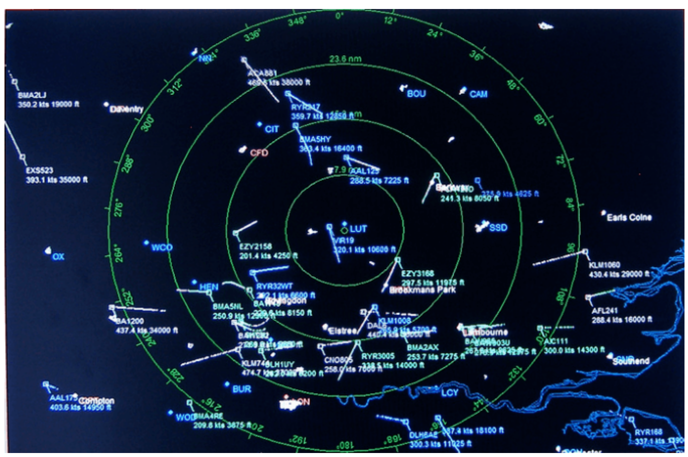


Figure 2. Complex data display and overlapping text.

https://www.pagerpower.com/news/radar-mitigation-frodsham-wind-farm/

Figure 3. Complex visualization with pop-up dialog box text.

In terms of effectiveness, the need to incorporate text into a complex visualization may have unintended consequences. This is especially true considering the relatively large display surface area required to ensure that visually displayed text be readable via the unaided human visual system. Either on the written page or electronic display media, minimum levels of brightness, contrast, character size, and spacing must be achieved for readability to be acceptable.[2] The value of visualization can be summed up with the classic axiom: “a picture is worth a thousand words,” but this ignores the question “what must a word be worth to justify occluding the picture?”

In some applications, there is often more information the interface designer wishes to make accessible than physical display surface area available to avoid unacceptable occlusion. Regardless of the display scale under consideration, when text presentation is deemed necessary, the effectiveness with which it can be included in the visualization is an important design challenge. If a large amount of text is displayed at once, such as inside a pop-up dialog box, significant regions of the “God’s Eye View” may be occluded, potentially occluding the view of a large number of monitored entities. When attempting to display text within a small area, a readability and/or accessibility cost may be incurred. The objective here is to design and evaluate a generalizable competing method for affording operator access to text-based information within operationally representative complex and dynamic visualizations. Concepts of interest share a simple design question: how can the amount of readable text-based information available to a visualization user be maximized while minimizing the display surface area required for its availability and accessibility? This is a driving motivation behind the design and evaluation of a “spatially-constrained” text access format.

Considering the physiology of the human visual system, it is the case that only a relatively small amount of displayed detail can be resolved during a single fixation. As presented by Cornsweet, the cone type photoreceptors of the human retina are exclusive to and densely packed into the fovea. Because of the density of the cones within the foveal area (approximately 140,000 cones per mm^2 within one degree of radial area at the center of retina), and the optical physics of the human eye, which focuses incoming light onto the fovea, fine detailed visual information may only be perceived within a small portion of the visual field.[3] Regardless of the physical display surface size, at any given distance from the display and under adequate luminance levels, the human visual system is able to resolve the highest level of information detail at just less than 4 degrees subtended visual angle. Detail falls off significantly and rapidly across the next 5 degrees.[4] To “see” the detail information content of a display area, the area must be scanned by a series of eye movements with associated sampling fixation points. The eye movement between the fixation points is called a saccade. The perceptual phenomena described here have implications regarding how humans read text as well as how the overall content of displayed information is visually sampled.

The focus of the present research is an empirical investigation of the relative merits between a pair of techniques operators may use to view and access text-based information in the context of operationally representative tasks involving the display of complex visual information. The research includes the conceptual development of an alternate text access technique as well as an operationally representative evaluation environment and task scenario. Human performance measurements, task accuracy and completion duration, are the dependent variables of interest. This effort includes the balance of experimental control for purposes of isolating any performance effects attributable to the independent variables of interest with the desire to generalize the findings.

METHOD

This study compared human performance when using the traditional “9-line” standardized targeting information display with a spatially-constrained display having various interaction elements within a secondary-task paradigm. Secondary task performance was expected to be degraded by occlusions due to text overlays as well as delays in accessing information from each presentation technique.

Experimental conditions and apparatus

Within the experimental protocol, the participants began by monitoring a ‘God’s-eye’ view display and reporting changes of presence, identification, and direction of movement of multiple entities within the display. Measures of accuracy and response time were used as performance metrics for the secondary task. As participants were asked to prioritize the primary task over the secondary task, accuracy and response time differences may not exist within performance of the primary task across the independent variables of interest. The design of this secondary task has been described elsewhere.[5] Therefore, performance and accuracy of the secondary task served as the basis of comparison to measure the relative effectiveness of the text display and access technique variants.

The primary task was initiated by a query presented to the participant at the top of the display to ask detailed information about a particular entity that required the use of the 9-line information. This task mimicked an operational task where it is necessary for the operator to confirm specific information contained within the 9-line data associated with an entity of interest. The text-based query contained the identification of the entity of interest, the 9-line category where the information is contained, and the specific data value to be confirmed. Figure 4 is an illustration of how a text probe query was presented to a participant.

When a query was presented to the participant, the participant’s task was to access the appropriate information for the identified entity of interest and confirm that the information associated with the entity is an accurate match (mouse over and select the “Yes (Y)” input) or that the information associated with the entity is not an accurate match (mouse over and select the “No (N)” input). A small input box was presented so the Yes/No input could be logged. When the cursor was hovered over one of the Y/N buttons, the button nearest the cursor tip was highlighted. Whatever button was highlighted when a mouse button was clicked was recorded and logged as associated input. Figure 4 includes an illustration of confirmation input selection Y/N buttons. The query report buttons were present anytime a valid query was present and are removed after an input was made or the query reached the timeout duration. The total number of positive and negative but accurate responses were balanced across the trial blocks while presentation query response type was randomized.

By explicit instruction, participants were requested to treat the primary task as their highest priority. Further, to avoid unwanted speed/accuracy tradeoffs, the instructions indicated priority of accuracy over speed within the query task. Participants were directed to perform the secondary task only to the extent there was spare capacity to do so.

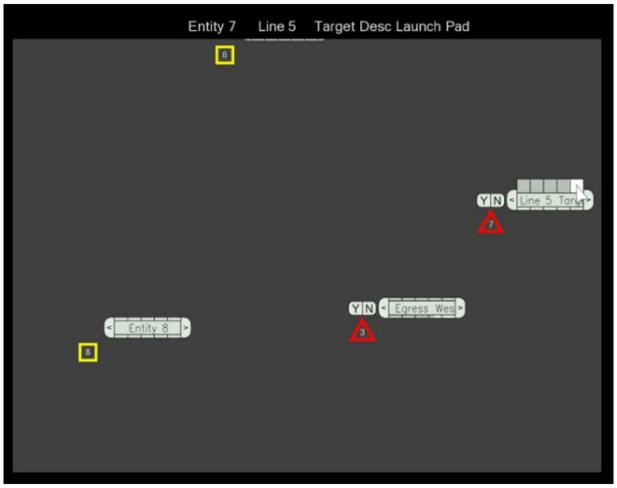


Figure 4. Example primary task query probe.

Baseline display condition (pop-up dialog box (DB) format)

The baseline condition consisted of a conventional pop-up dialog box. Each active entity in the wide God's-eye display had an associated ID data tag. The data tag was a fixed size and accommodated approximately 13 characters. The tag box (ID tag) size was based on the human ability to perceive 12-14 characters in a single fixation with acceptable accuracy.[1,6,7,8] Figure 5 (as presented by Öquist), demonstrates the foveal and parafoveal accuracy approximation within a single human eye fixation.[1] For this investigation, it was desirable that the entity label be perceived with a single fixation and that the alphanumeric characters of the ID tag require little to no saccadic eye movement to be read.

The 9-line display presents vital data elements to support a targeting or sensor data collection activity. This display identifies the acting agent or performer of the mission element. For C^{3}I display purposes, this item is typically a call sign label. The additional 9-line information underlying allows command and control elements to understand the assignment and intent of the acting agent.[9] For clarity and ease of use, the items contained in the 9-line are standardized to afford monitoring and improve situation awareness at the single entity level. Also, as mission changes occur and new assignments are passed to entities of interest, the 9-line information forms the basis of a closed communication loop between command elements and actors. The items in the standard 9-line sequence include: call sign, 1) initial point, 2) heading, 3) distance, 4) target elevation, 5) target description, 6) target location, 7) marker type, 8) location of friendly forces, 9) egress, and *10) remarks*. To access the 9-line information associated with the entity, a participant selected the ID tag via placing the cursor on the entity and performing a left click.

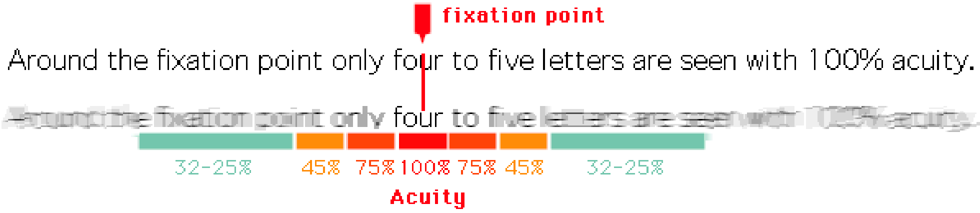


Figure 5. Illustration of single fixation foveal and parafoveal perceptual accuracy.

The vertical size of the 9-line dialog box varies depending on the amount of text required of the 9-line assignment. Text is wrapped to fit within a window size as defined by the often used "page" format containing a width limit of 50 characters.[10] Figure 6 shows a rendering of a completed 9-line page pop-up dialog box. The window stays open until the cursor is located within the window and another left-click is input via the mouse controller. This interaction feature enables the user to keep multiple windows open for multiple entities as long as desired. It is expected that this interaction mechanization will be familiar and intuitive for users and will require little or no usability training.

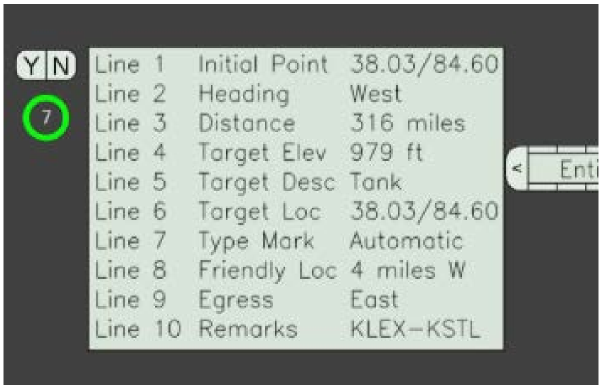


Figure 6. Example entity with pop-up dialog box containing 9-line information.

Spatially-constrained display

The baseline display was compared to a spatially-constrained display format. A review of the previous research revealed the following considerations for this display.

1) A single fixation between saccadic eye movements translates to an identification level resolution of about 10-13 letter spaces of text for a typical font size and viewing distance.[5] This value defines an upper boundary for the size of a single fixation display.
2) Research showing the potential for rapid serial visual presentation (RSVP) was consistent with this value when replicated on electronic displays (RSVP described below).[11,12]
3) Studies performed to "tune" text presentation for human reading behavior, such as saccadic eye movement, showed performance benefits.[13,14,15] This indicates that a similar feature should be included when text is in full sentence form.
4) The addition of progress monitoring and direct user control of spatially-constrained text presentation showed promise toward both enhancing performance and improving user acceptance.[1,13,16] This motivates the use of interactive elements for the present research.
5) Rapid serial presentation of changing system values showed performance benefits compared to natural language text presentation.[17,18,19] The value-like type of information presentation is more similar to the present research text content versus the presentation of complete sentences used in reading research.

Figure 7 shows a rendering of a proposed spatially-constrained entity tag with user interaction symbols. Similar to the baseline condition, the drawing area of the data tag accommodated approximately 13 characters at any one time.

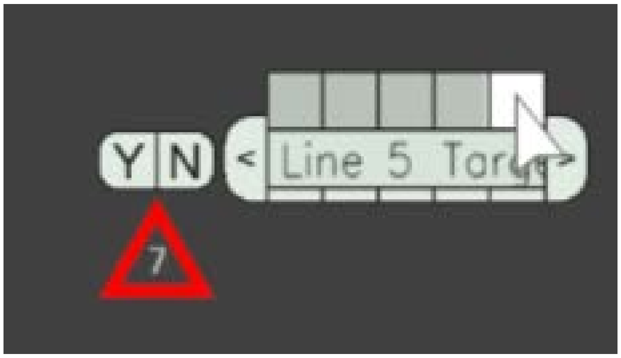


Figure 7. Example of a spatially-constrained window and function keys. Entity shape includes the addition of an ID number.

Forster performed experiments to determine if flashing the words of a sentence at a fixed position could demonstrate humans' ability to read information without the need for saccades.[12] The presentation technique was called rapid serial visual presentation (RSVP). Although Forster's stated objective was to determine if RSVP could be used to detect sentence complexity, this series of studies simultaneously provided a methodology for evaluating RSVP reading efficiency.[12] Words were presented to participants at a rate of 16 words a second (62.5 ms each) and each word was projected onto the retina at approximately the same area. Broadbent and Broadbent used a variant of the RSVP method to investigate a phenomenon where target word identification performance suffered from interference under conditions when the target was defined in some way that did not specify its full identity.[11] This is analogous to a task where a user is interested in knowing that some change in the information has occurred and is further interested in being able to identify the nature of that change. Castelhano and Muter described research performed to investigate the best presentation techniques for the display of text on small electronic screens.[13] Some obvious analogies can be drawn between display via small screen and the display of text within a small drawing area regardless of the overall display screen size. The study is a continuation of work previously performed by Kang and Muter and Rahman and Muter.[14,15] Here, RSVP was compared to several other presentation techniques such as a moving window display, right to left horizontal scrolling, line-stepping, and sentence by sentence presentation. It is stated that the efficiency gain to be realized by RSVP was likely due to time savings related to the reduction of saccadic eye movements (and associated cognitive load) compared to that required during reading a conventional page layout. While RSVP has application potential, it is typically disliked by readers compared to more conventional reading methods.[13,14,20]

From all the literature presented thus far, an interesting observation can be made. In the years since the development and evaluation of the research presented above, the use of small screen technology has been widely embraced, and its dependence on text-based information continues with a more or less conventional page-based presentation. Although the RSVP techniques demonstrated performance advantage, there was little evidence that any practical application was found beyond some value as a research tool. Of course some dynamic text presentation techniques are quite common. Öquist performed a series of experiments with the aim of finding the best way to present text-based information on screen sizes appropriate for mobile devices such as personal digital assistants (PDAs) and mobile telephones (presently, a combination of these technologies exists as the "smart phone").[1] Scrolling, leading, paging techniques, and RSVP were investigated via the use of eye tracking measurement and task loading index ratings.

Based on Öquist's reported results, it is hypothesized that user interaction with the data may be a key attribute of a spatially-constrained display, making it possible to access specific elements of information on demand.[1] Therefore, the spatially-constrained text access display will enable significantly more interaction than the baseline display. To enable this interaction, the upper and lower bezel of the spatially-constrained window was segmented into 10 small selectable "soft keys". Hovering the mouse cursor over the soft keys results in their graphical expansion so that selection is easier than the original size of the keys. This design is conceptually similar to interaction with a multi-function display. Each of the bezel keys correspond with the respective 9-line element. The 10th key is reserved for any "remarks" information that the entity 9-line may include. When the participant selects one of the numbered keys on the bezel (via a mouse-driven cursor left-click), the corresponding 9-line text is presented dynamically as an RSVP.
The RSVP will toggle between run or pause when the user left clicks within the area of the presentation window. When running, the 9-line text will begin with the first category line, proceed in sequence to the end of the 9-line information and repeat the cycle until some other mode is selected. Current cycle duration is set at 20 s from beginning to end. This translates approximately to a 6-8 character/second presentation rate. Left clicking in the window pauses the text within the drawing area, presenting a static string of text that is visible within the dialog box at the instant the left-click selection is made. If the user takes no other action for 500 milliseconds, the mode automatically switches to a "stop" state and only the top level entity label text is presented within the presentation window.

Fast forward and reverse are enabled by left clicking on the chevron symbols. A fast-forward action during run or direct selection mode will skip the text presentation to the beginning of the next line in the 9-line sequence. Depending on the initial mode prior to this action, the respective presentation continues from the new start point. To activate a fast forward, the user hovers the mouse-driven cursor over the fast-forward soft key and left clicks the mouse input button. Left clicking the left-pointing chevron symbol begins the 9-line sequence again from the beginning of the present text line if selected toward the last 2/3 of that line. If selected during the beginning of the text line (first 1/3), the reversion action snaps to the previous line and begins to run.

Feedback regarding the current location within the 9-line sequence is provided by changing the amount of fill within each category indicator. For instance, a half-filled key indicates half of that line of text has been presented and half is yet to be presented. Line presentation completion corresponds closely with the key being completely filled. When presentation switches to the next line, the fill animation begins again for that line key. This functionality provides a means for the user to become quickly oriented to the location of the displayed text within the 9-line. Also, this approach provides a smooth and unobtrusive cycle progress tracking capability.

It is also possible that the motion of the element may interact with either display technique, but particularly with the RSVP technique, making it useful for the spatially-constrained window to remain stationary on the display while selected. For this reason, the experiment contained an additional variable, wherein the text access window followed the motion of the entity or was stationary once displayed.

Presentation window entity motion dependence

When the text presentation window is selected, the way it (the window) behaves after that may have a significant effect on performance. In this condition the text presentation window behaves as if it is "attached" to the entity as it moves across the display. The location of the text presentation window appears to the upper right of the entity and is tethered to the entity in a way that it will move within the x and y dimensions of the allowable scenario area. For this level the interaction mechanization variable is called Dynamic Dependence (DD). A second level of is called Static Independence (SI). In this condition, once the text presentation window is selected by the user, it remains stationary within the display regardless of entity motion.

Other important behavioral attributes

It is likely a requirement that some ID label stay with each entity even though it moves away from the window displaying the 9-line information within the DD condition. This is accomplished by superimposing an ID number within each entity shape as shown in Figure 7. Regardless of the condition, when selected by the participant, the 9-line information associated with the entity of interest becomes active and the presentation continues until the entity is "de-selected" by the user. This functionality allows the information associated with multiple entities to be available simultaneously.

Relative to each other, all the entities, data tags, and text access windows are opaque. To assign occlusion, a layering approach is used based on the order in which entity interaction was sequenced. The most recent interaction brings those objects (set combinations of entities and their associated data tags and activated text access windows) to the top layer and sends each other object one layer back from its previous layering sequence. Similarly, order of entity presence determines layer position in the absence of user interaction. The most recent object with the most recent interaction is at the top or "front" layer. The occlusion sequence described above is the same regardless of whether the text access interaction mechanization is DD or SI.

All conditions were displayed on a standard 30" LCD monitor. Participants were seated approximately 60 cm from the display.

A conventional computer equipped with standard keyboard and cursor input device (mouse) was used to generate the stimulus display and used for participant response input/questionnaire responses. Questionnaire responses were also collected via pen and paper recording. Software code development for this study was performed in-house. The code includes generation of all the visual elements, task mechanization, performance data collection routines, raw data recording, and any automatic data reduction.

Participants

To date, 9 participants have completed a full set of data collection for the experiment. Each participant took part in all conditions of the study. Participants were between the ages of 18 to 60 and self-reported normal or corrected-to-normal visual acuity and color vision. Participation duration including introduction/consent, instructions, demographics/vision screening, data-collection, and rest breaks was approximately 320 minutes. Each experimental condition lasted 10 minutes.

Experimental design

The study employed a 2 by 2 by 4 full-factorial within-subjects experimental design. The experiment included 2 levels of text access format (DB and RSVP) and two levels of interaction mechanization (DD and SI). The final factor was four replications intended to add variation to the participants' experience.[5] Order effect confounding for the repeated measures variables was mitigated by counterbalancing. To analyze effectiveness given the interdependency between response time and response accuracy, an analysis of the independent variables using the response duration measures while holding the accuracy variable constant was utilized. The question being, what is the effect on response time when accuracy is always 100% correct?

Nine (9)-line information values

The 9-line data generated for the study was selected so that it meets specific criteria. The data values were simple but realistic and made to appear to be what would be expected of operationally relevant information as if the entities were individual aircraft with tactical assignments. The data were formed of latitude, longitude, and elevation coordinates, heading directions, distances, target types, locations, etc. The data were selected to be easily and quickly understood as well as easily held within the limitations of working memory by minimally trained naïve participants. No specific skill set is required to perform the tasks. Further, the number of individual text characters within a query probe data value is designed to be accurately perceived by the participant within the limitations of a single foveal fixation. This is based on the normal human ability to perceive 12-14 characters in a single fixation with acceptable accuracy.[1,6,7,8]

The data given in Table 3 are the basis for both the query input and the 9-line content associated with the specific entities to be displayed via the two different formats and two different levels of the interaction mechanization variable. For a positive or "Yes" confirmation, the query and entity 9-line data items match. For a negative or "No" confirmation response, the query item represents the correct category of information in terms of line number and description but selection logic included the limitation that the data item origin entity and the query entity could not be the same thus, the data items were the appropriate type but different values. As a result, participants were not able to determine a correct query response by a recognizing that the query data item was the wrong category of information for the query line number and description. A correct response included confirmation of matching data items when the query entity and the selected entity are also the same. Similarly, a correct response was recorded when the participant selected "No" when the query entity were the same but the data items did not match.

Query event control

Query events occurred according to a controlled schedule. Per 10 minute (600 s) data collection session (each a single trial), there were a total of 22 queries (approximately one every 25 s). The first 25 s period of a trial did not include a query so that enough time was allotted to allow the scenario to be fully populated with secondary task entities prior to the initialization of the first query task. Additionally, a "blank period" of 25 s occurred at the 300 s point of the trial to balance the number of query events in the first and second halves of each trial. The initialization of each query was randomized within the first 5 s of each query period. This resulted in less predictable query timing. The remaining 20 s of each period is allotted for completion of the query task. If the task was not completed within that 20 s window, the query task timed-out and was recorded as such.

Table 3 Nine (9)-line data content.

Line 1	**Line 2**	**Line 3**	**Line 4**	**Line 5**	**Line 6**	**Line 7**	**Line 8**	**Line 9**	**Line 10**
Initial Point	Heading	Distance	Target Elev	Target Desc	Target Loc	Type Mark	Friendly Loc	Egress	Remarks
38.37/81.59	North	163nm	1202ft	Runway	40.49/80.23	Visual	8nm NW	South	KCRW-KPIT
42.40/83.01	Northeast	521nm	626ft	Fuel Tank	42.40/83.01	Laser	5nm NE	Southwest	KDET-KBUF
39.04/84.66	East	413nm	896ft	Hangar	39.04/84.66	Infrared	8nm SW	West	KCVG-KRIC
36.19/95.88	Southeast	230nm	677ft	Landing Pad	36.19/95.88	Beacon	6nm NE	Northwest	KTUL-KLIT
36.08/98.15	South	239nm	2181ft	Bunker	36.08/98.27	Talk on	5nm S	North	KLAS-KNYL
33.63/84.42	Southwest	147nm	1026ft	Tower	33.63/84.42	Radio	7nm S	Northeast	KATL-KMGM
38.03/84.60	West	316nm	979ft	Tank	38.03/84.60	Auto	4nm W	East	KLEX-KSTL
39.90/84.21	Northwest	241nm	1009ft	Launch Pad	39.90/84.21	Code	9nm N	Southeast	KDAY-KORD

Primary task dependent measures

The primary task dependent measures of interest (within each cell of the balanced test matrix) included response accuracy and response time. The accuracy measure for the primary task was simply a correct response (confirmation that the query and target data values match when they are supposed to and confirmation that query and target do not match when they are not supposed to), or an incorrect response (indicating that the data match when they in fact do not or indicating the data do not match when in fact they do). These responses are recorded as hits, correct rejections, false selections, and false rejections.

The following primary task dependent measures were recorded for purposes of findings analysis and findings interpretation:
1) Accuracy measures: correct, incorrect, and incorrect type.
2) Response duration measures: total response time per query task, elapsed time from query presentation to entity selection.
3) Timeout occurrences.

4) Subjective Workload: recoded at the session and trial level (text access format with entity dependence).
5) Text access technique subjective preference recorded after the participant has experienced all of text access formats.

Secondary task dependent measures

The secondary task dependent measures were largely focused on accuracy and response duration. The response accuracy measures for the secondary task were based on proper selection by the participant if the presence, identity, or direction of movement for an entity changed. Errors are defined several ways: 1) failure to indicate whether the entity changed entry presence, identity, or direction within 10 s of the change; 2) selection of an entity undergoing change but incorrect selection of the type of entity, or 3) selection and input for an entity not associated with a preceding change.

Task training and feedback

After completing a written and verbal description of the experimental tasks, participants were trained on the primary, secondary, and combined tasks until the effects of learning on performance became asymptotic. Participants were trained first on the secondary task until performance was nearly error free and workload was acceptable. Participants were instructed to maintain a continuous visual scan pattern so that change events among the entities were not missed. Next, participants were familiarized with the primary task by performing it in isolation of the secondary task. Participants experienced the primary task with each of the independent variable levels. Once participants were comfortable with the primary task and performance stabilized, the primary and secondary tasks were combined to train simultaneous performance of both the tasks. It was reiterated that the participants should prioritize performance on the primary task and, within both tasks, prioritize accuracy performance over speed. Participants were reminded to perform the secondary task to the extent possible given any spare capacity to do so.

During the training sessions, participants were provided feedback to help inform them firsthand of their accuracy in the primary, secondary, and combined tasks. Training performance and feedback were monitored by the experimenter so that errors were pointed out and tips for optimal performance of the tasks were communicated and standardized across participants. The objective was for all participants to perform every task as similarly as practical so that their performance is representative of trained operators. This process also acted to ensure that all participants have the motor dexterity and skill to perform the experimental tasks.

Subjective feedback data collection

Additional measures of interest were subjective. It is important to collect subjective workload measures at appropriate points within the overall test matrix to give a global measure of participants' perceived level of effort (based on spare capacity). Subjective preference data were also be collected to analyze consistency or lack of consistency with the other measures. Workload was reported using a Bedford Workload Scale decision tree procedure.[21] The workload rating scale is a fairly self-explanatory procedure where participants report subjective task workload based on consideration of spare capacity to perform additional tasks with regard to the presently performed tasks. A practice data collection (task training) session included use of the scale.

After completion of all data collection sessions, participants completed a rating questionnaire designed to record their subjective preferences for the combinations of the independent variables related to the performance of the primary task. Additionally, participants indicated how well they thought they were able to perform the primary task using the combinations of text access formats and interaction mechanizations.

RESULTS

This analysis is based on a partial dataset (data collection was not complete at the date of this publication) collected from 9 subjects. The analysis model is a 2 by 2 by 4 within-subjects Analysis of Variance (ANOVA) performed separately for the primary and secondary tasks for each dependent variable: Task Performance Accuracy and task completion Response Duration for Correct Responses. The first level independent variable factor is a comparison between the text access formats (Baseline (pop-up dialog box and RSVP). The second factor is comparison of interaction mechanization (DD and SI). The third factor is replication.

Primary task

For Task Performance Accuracy, all effects had large p-values (Access Format: p = 0.2824; Interaction Mechanization: p = 0.4289; Replication: p = 0.7294) thus no further investigation was conducted. For Response Duration (correct responses), Access Format and Interaction produced very small p-values (Access Format: p = 7.834 e-10; Interaction Mechanization: p = 1.690e-05) and the effect sizes (Cohen's d) are also large or greater (Access Format: d = 1.155; Interaction Mechanization: d = 0.6350). Thus, there is evidence to suggest that Format and Interaction Mechanization have an effect on Average Response Duration when responses are correct. No other main effect or interaction is evidenced as influencing this measure. The nature of the Access Display Format effect is shown in Figure 8 and similarly, Interaction Mechanization is shown in Figure 9.

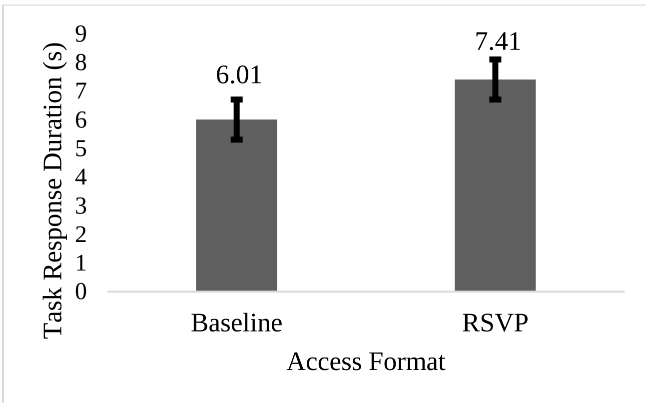


Figure 8. Primary task average response duration as a function of text access format.

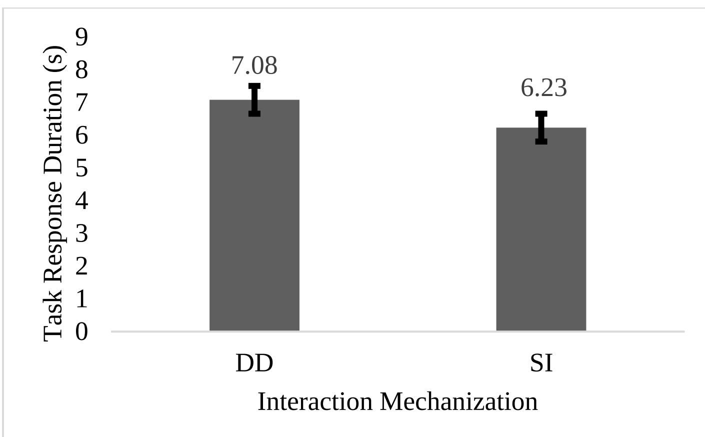


Figure 9. Primary task response duration as a function of interaction mechanization.

Secondary task

For Task Performance Accuracy, the effects had fairly large p-values (Access Format: p = 0.0484; Interaction Mechanization p = 0.0944; Replication: p = 0.9537) and the corresponding effect sizes are small to medium (Access Format: d = 0.3678; Interaction Mechanization: d = 0.2872). No other main effect or interactions showed evidence of having an effect on Accuracy in the secondary task. For Average Response Duration (correct responses), Access Format and Interaction Mechanization resulted in small to very small p-values (Access Format: p = 0.0358; Access Format: p = 0.0002) and the effect sizes are shown to be small to medium (Access Format: d = 0.3723; Interaction Mechanization: d = 0.6967). Thus, there may be evidence to suggest that Access Format and Interaction Mechanization have a slight effect on Average Response Duration in the secondary task. No other main effect or interaction is evidenced as influencing this measure. The nature of the effect of Access Display Format is shown in Figure 10 and similarly, Interaction Mechanization is shown in Figure 11.

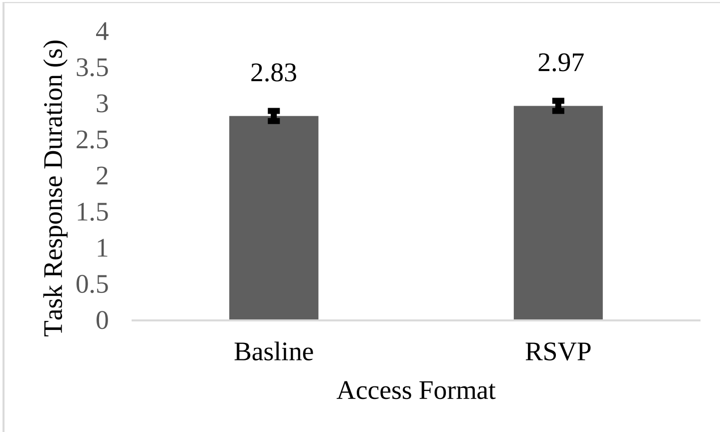


Figure 10. Secondary task response duration as a function of text access format.

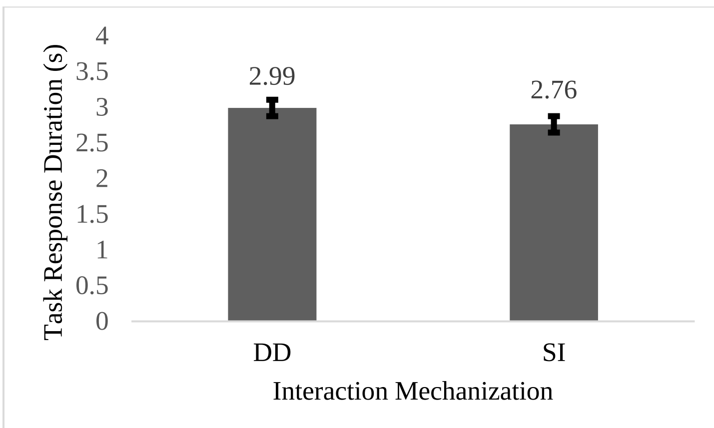


Figure 11. Secondary task response duration as a function of interaction mechanization.

<u>Subjective feedback</u>

For Standardized Workload Rating (utilizing the Bedford Workload Scale), all main effects reported large p-values (Access Format: $p = 0.1271$, mean rating = 5.8/10; Interaction Mechanization: $p = 0.3347$, mean rating = 5.8/10; Replication: $p = 0.9845$). Effect sizes are also small (Access Format: $d = 0.2859$; Interaction Mechanization: $d = 0.2037$). There is no evidence that the independent variables had a meaningful effect on reported subjective workload although, overall reported workload was fairly high.

In a direct comparison of the text access formats in terms of preference, for the data collected to-date, only one subject out of a total of nine responses favored RSVP over the baseline format.

DISCUSSION

The objective of the present study was to evaluate a novel approach for affording access to relatively large amounts of text information while using a comparatively small amount of display real estate. The main experimental comparison between a baseline text access format in the form of a pop-up dialog box and a miniature multi-function display used a secondary-task paradigm to collect quantitative performance measures. Participants were instructed to prioritize the primary task (accessing text content via the different display formats to address an information query) over the secondary task (event monitoring and reporting). Similarly, participants were instructed to prioritize task accuracy over task completion speed. A second independent variable manipulation included a text access box which moved with the entity of interest (dynamic dependent) or whether the access format was fixed in screen coordinates (static independent). Similar quantitative measures were used to evaluate these interaction mechanizations.

The combination of display features which most closely represent the conventional means of presenting detailed text information is the baseline format and the static independent interaction mechanization (e.g., Figure 3) versus the spatially-constrained RSVP format and with dynamic dependence. The RSVP format requires that the user interact directly with the features of the text presentation. The intended advantage of the spatially-constrained text access design is that it requires relatively small amounts of display real estate and thus it may occlude and interact less with "big picture" tasks such as entity monitoring. According to the results, there was not clear advantage shown for the RSVP text access format or the dynamic dependent interaction mechanization. On the other hand, where task accuracy was measured, there was not a decisive disadvantage either. Given the prioritization of the primary task over the secondary task, the evaluation was designed to detect performance disadvantages as a result of occlusion of the reportable secondary task events by pop-up dialog box. The results indicate that if there was any advantage at all, the baseline format was favored both in terms of accuracy and completion duration. It is possible that the number of entities in the monitoring task and the speed with which the primary task could be completed did not lend itself to a robustly measurable difference due to occlusion. An analysis of the physical area used by the display formats over time may offer some insight into the potential for the display formats to differentially affect performance.

Differences among the independent variables where statistically supported when task completion duration for correct responses was compared. For the primary task, both the baseline format and the static interaction mechanization showed some advantage. This was not surprising given the additional steps required for the subject to access the RSVP text presentation. On average, the duration advantages detected for task completion did not exceed 1.5 seconds when the response was correct. Given the application, these differences may or may not represent a burden compared to the potential features represented by the spatially-constrained text access format. It is notable that no subjective workload difference was detectable among the variables. A manipulation of the text presentation dynamics (RSVP versus others) my help tease out whether or not there is room to optimize the efficiency of the spatially-constrained concept.

Past research has shown that RVSP performance is superior compared to some other text presentation techniques but in general, subjects tend to dislike RSVP.[13,14,15,20] This finding seems to be replicated here. Again, other presentation techniques may remedy this.

Given the level of novelty represented by the spatially-constrained text access concept and the lack of clear performance disadvantage when compared to a well-known conventional electronic display information access interface (human-system interaction), the new approach shows promise. An effort to develop an optimized format and associated interaction features seems warranted as this technique is likely to have greater advantages on highly cluttered displays.

AKNOWLEGMENTS

The authors appreciate the helpful experimental design, analysis, and concept suggestion inputs by Drs. Christine Schubert Kabban and John Colombi at the Air Force Institute of Technology. The authors would like to thank Eric Heft of the Air Force Research Laboratory's 711th Human Performance Wing for his expert coding of this experiment and the data collection routines. Our appreciation to Lauren Menke of the 711th Human Performance Wing for her assistance during subject scheduling and data collection. Also, the authors appreciate the time and effort expended by the subjects who volunteered to participate in this study.

PRIMARY AUTHOR BIO

Eric E. Geiselman is a Senior Engineering Research Psychologist at the U.S. Air Force Research Laboratory's 711th Human Performance Wing, in the Human Effectiveness Directorate. His research interest is concentrated in the area of visual display operator interface design and evaluation. Mr. Geiselman spent 8 years as an airline pilot and eventually became a Crew

Resource Management instructor. Mr. Geiselman holds an M.A. in Experimental Psychology from the University of Dayton and is pursuing a Doctorate in Systems Engineering at the Air Force Institute of Technology.

REFERENCES

[1] Öquist, G. (2006). *Evaluating readability on mobile devices*. Acta Universitatis Upsaliensis. Studia Lingustica Upsaliensia, Uppsala. ISBN 91-554-6745-8.

[2] Kruk, R. S., & Muter P. (1984), Reading of continuous text on video screens. *Human Factors*, 26(3), 339-345.

[3] Cornsweet, T. N. (1970). *Visual perception*. Academic Press, New York, NY.

[4] Le Grand, Y. (1975). Measurement of the visual stimulus. In *Handbook of perception volume V seeing* Eds. Carterette, E. C. & Friedman, M. P. Academic Press, New York, NY.

[5] Geiselman, E.E, Heft, E., & Miller, M.E. (in review). Development and Validation of a Secondary Task Environment for Assessing Visual-Psychomotor Tasks. *Theoretical Issues in Ergonomics Science*: Taylor & Francis.

[6] Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 374-422.

[7] Just M. A., & Carpenter P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychological Review*, 87(4), 329-354.

[8] Robeck, M. C. & Wallace, R. R. (1990). *The psychology of reading: an interdisciplinary approach, 2nd edition*. Hillsdale, NJ: Lawrence Erlbaum Associates.

[9] Air Land Sea Application Center. (1997). J-FIRE multiservice procedures for the joint application of firepower. TRADOC-MCCDC-NDC-AFDC Air Land Sea Application (ALSA).

[10] Piolat, A., Roussey, J., & Thunin, O. (1997). Effects of screen presentation on text reading and revising. *The International Journal on Human-Computer Studies*, 47, 565-589.

[11] Broadbent D. E., & Broadbent, M. P. (1987). From detection to identification: Response to multiple targets in rapid serial visual presentation. *Perception and Psychophysics*, 42(2), 105-113.

[12] Forster, K. I. (1970). Visual perception of rapidly presented word sequences of varying complexity. *Perception and Psychophysics*, 8(4).

[13] Castelhano, M. S., & Muter, P. (2001). Optimizing the reading of electronic text using rapid serial visual presentation. *Behavior & Information technology*, 20(4), 237-247.

[14] Kang, T. J., & Muter, P. (1989). Reading dynamically displayed text. *Behaviour & information technology*, *8*(1), 33-42.

[15] Rahman, T. and Muter, P. (1999). Designing an interface to optimize reading with small display windows. *Human Factors*, 41(1), 106-117.

[16] Muter, P. (1996). Interface design and optimization of reading of continuous text. In *Cognitive aspects of electronic text processing* (pp. 161-180). Ablex, Norwood, NJ.

[17] Matin, E. & Boff, K. R. (1988) Information transfer rate with serial and simultaneous visual display formats. *Human Factors*, 30(2), 171-181.

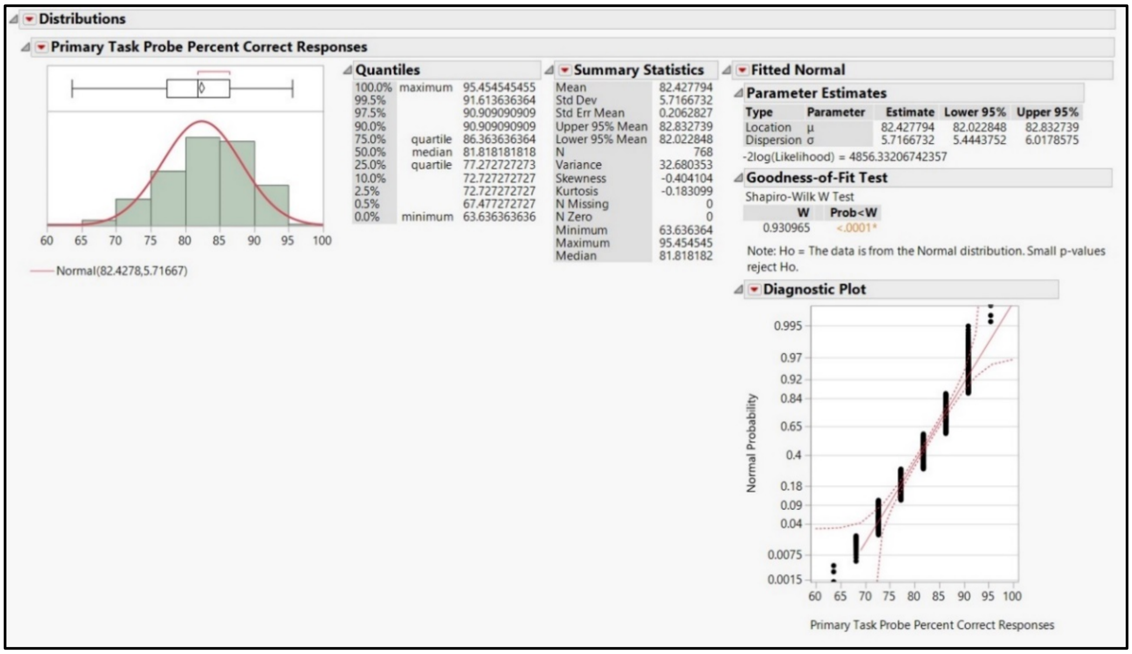
[18] Osgood, S. S., Boff, K. R., & Donovan, R. S. (1988). Rapid communication display technology efficiency in a multi-task environment. *Proceedings of the Human Factors Society 32nd Annual Meeting*, 1395-1399.

[19] Payne, D.G., & Lang, V.A. (1991). Visual monitoring with spatially versus temporally distributed displays. *Human Factors* 33(4), 443-458.

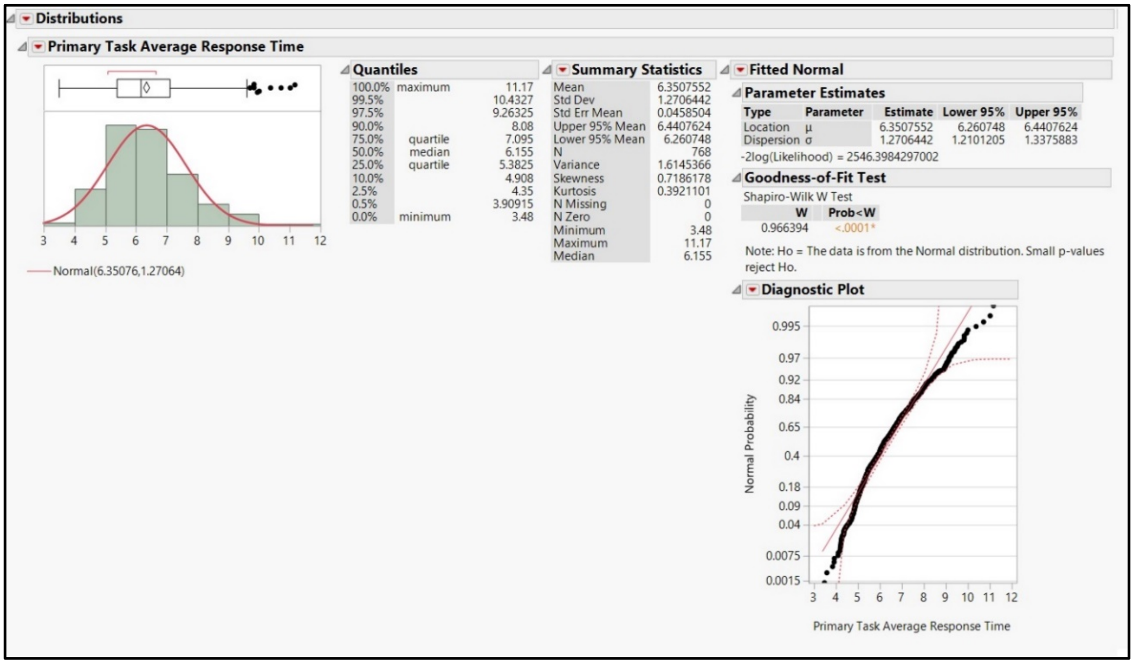
[20] Muter, P. (1996). Interface design and optimization of reading of continuous text. In *Cognitive aspects of electronic text processing* (pp. 161-180). Ablex, Norwood, NJ.

[21] Roscoe, A. H. & Ellis G. A. (1990). A subjective Rating Scale of Assessing Pilot Workload in Flight: A Decade of Practical Use, A technical Report, Procurement Executive, Ministry of Defence Farnborough, Hampshire.

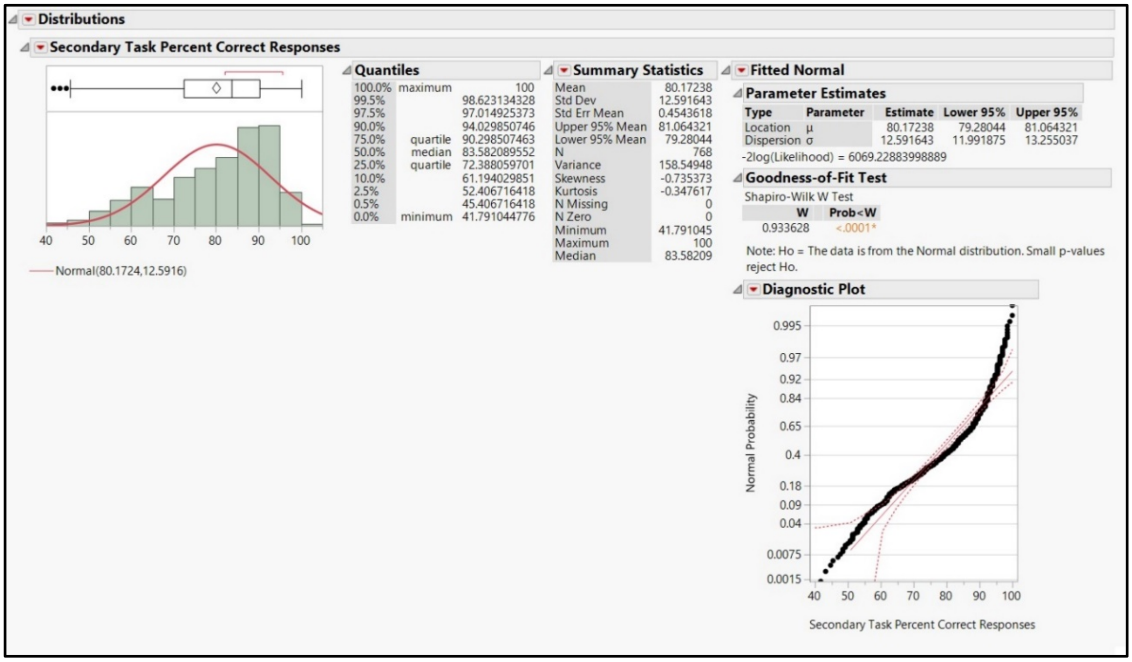
Appendix F: Selected Statistics Package Output



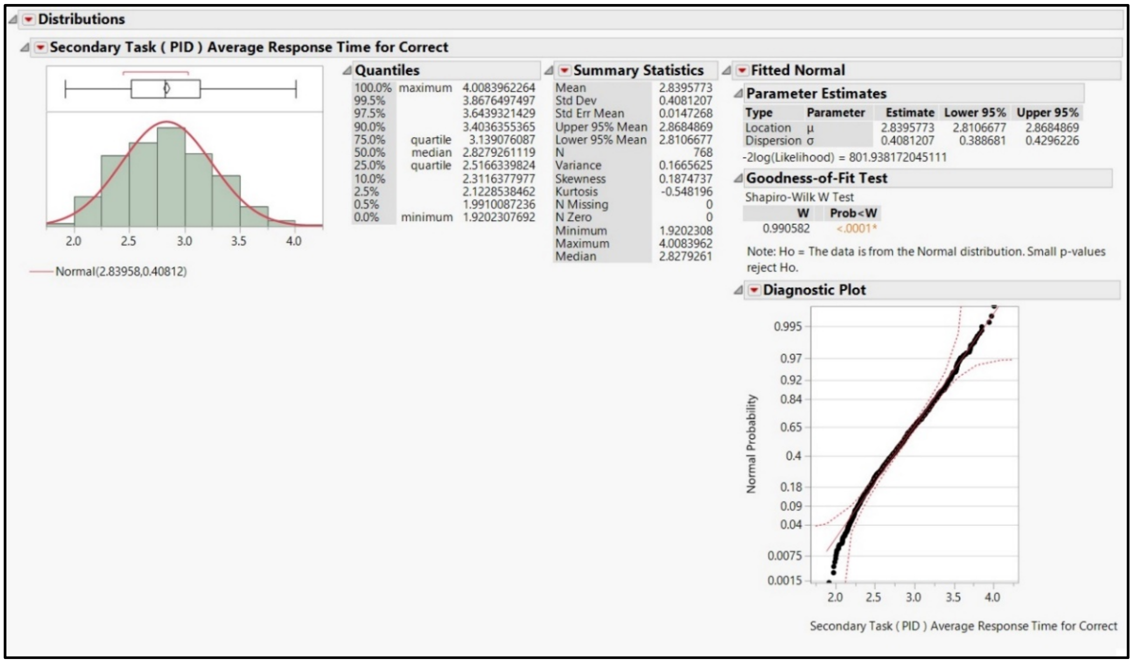
Primary task mean percent correct responses descriptive statistics and distribution.



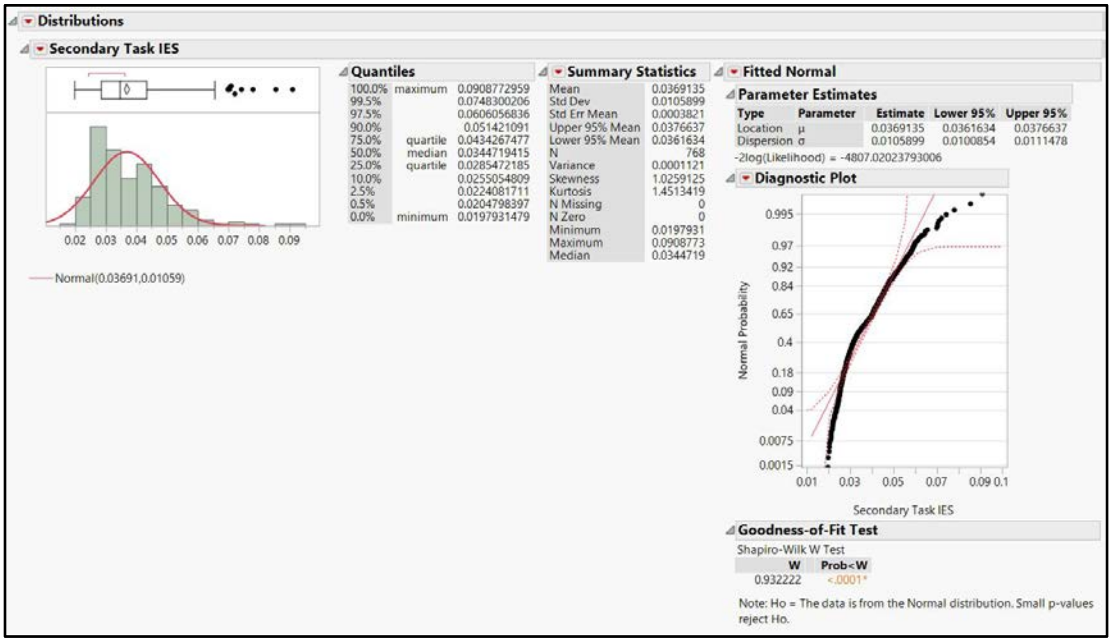
Primary task mean response duration for correct responses: descriptive statistics and distribution.



Secondary task mean percent correct responses: descriptive statistics and distribution.



Secondary task mean response duration for correct responses: Descriptive statistics and distribution



Secondary task inverse efficiency score: descriptive statistics and distribution.

Appendix G: Experiment 1 Additional Analysis

This appendix presents an additional analysis that was performed on the Experiment 1 data to augment the content of the published manuscript (Geiselman, Heft, & Miller, 2020). This analysis was performed to provide a better visualization of the interaction of event rate and entity number for the 30 second moving mean error measure (F (3.2, 11.8) = 8.6, MSE = 6.2, $p \leq .000$, $\varepsilon^2 = .49$) presented here as Figure G1.

Figure G2 shows entity and event rate mean error frequency within a 2 minute period plotted as a line graph to better characterize the nature of the interaction. From the figure it can be appreciated at the lowest event rate bin (2) that the mean error frequency for the 3 and 4 entity variable levels are superimposed and separated from the accuracy performance resulting from the 5 entity level. Comparisons were performed using Fisher's Least Significant Difference post-hoc tests ($\alpha = .05$). The star shapes in figure G2 indicate a point that is significantly different from the other two points within each event rate bin (event rate interval). A possible reason that performance was relatively good at the slowest event rate bin for the 4 entity level is that the random seed used to govern the entity behavior allowed the actual entity number under that set of conditions to remain lower for a longer period of time than desired. Since a single random seed was used throughout the experiment, the effect of an unwanted entity reduction during that set of conditions would have been persistent and propagated across data collection for all of the participants. Since the slow event rate bin produced very little error in general with low workload, this condition was of little experimental interest. Therefore, the negative impact of the potential confound was minimal.

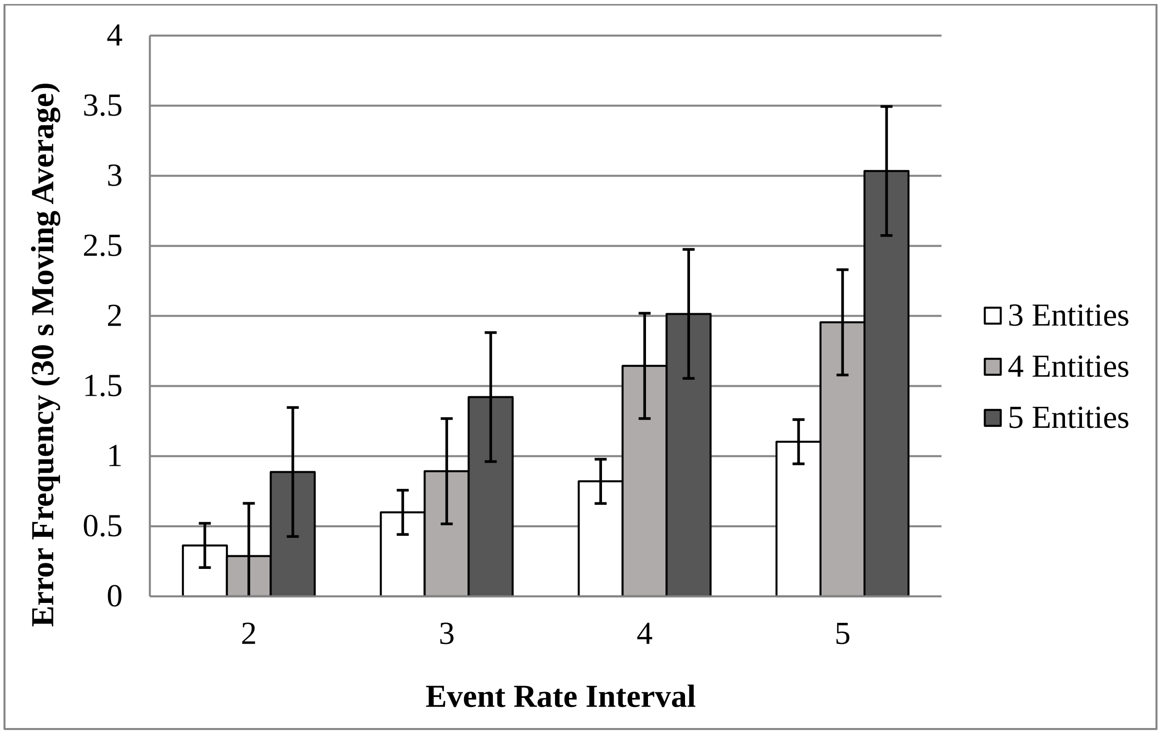


Figure G31. Mean error frequency within collapsed 2 minute event rate intervals across all participants.

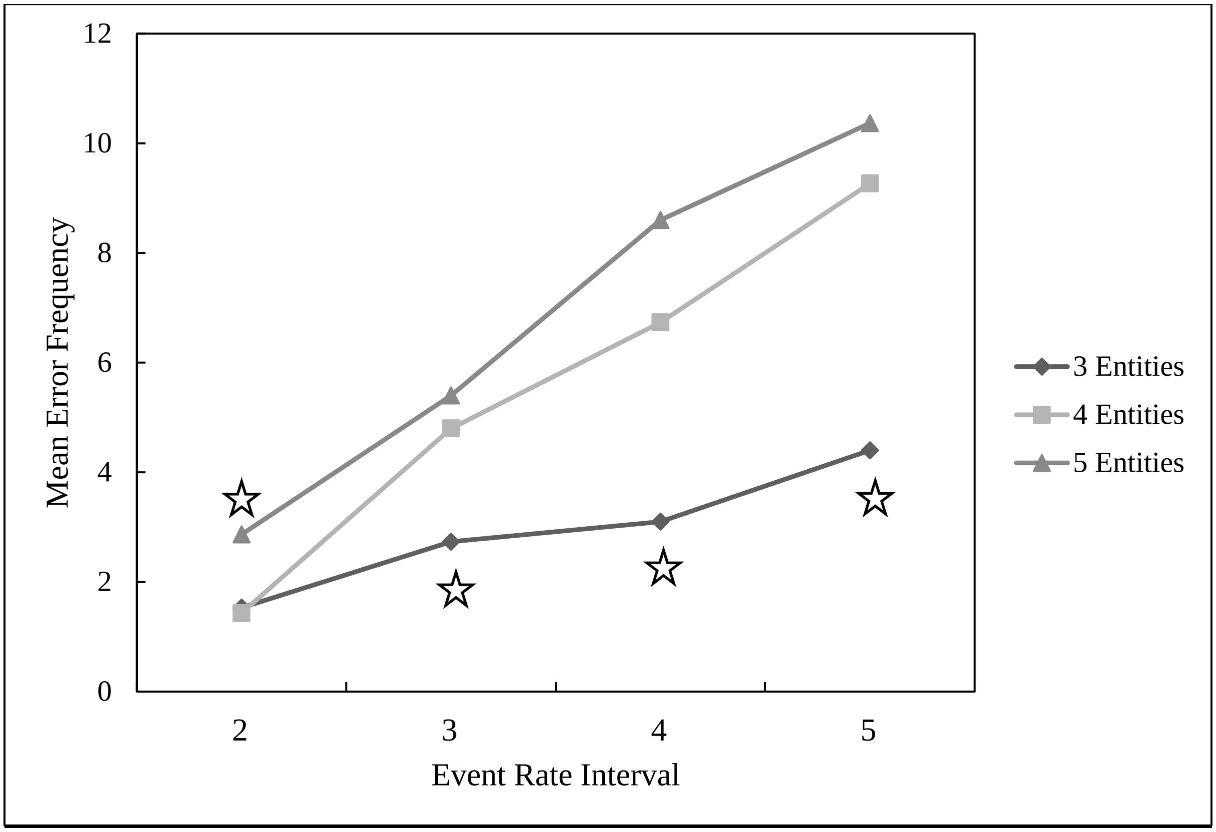


Figure G32. Mean error frequency as a function of event rate and entity number as a line graph.

When an average of 3 entities is maintained across all four event rate bins, it is evident also from Figure G2 that performance decayed less rapidly than was the case with either the 4 or 5 entity conditions. However, the slopes of the functions for the 4 and 5 entity conditions were approximately equal.

A complete set of post-hoc comparisons were performed across each of the 12 event rate (4 levels) and entity number (three levels) combinations. Each of the combined levels were compared to each of the remaining levels for the mean error frequency dependent measure. Figure G3 presents the plotted data set with the 12 comparison levels labeled (1 – 12) and their associated error frequency means.

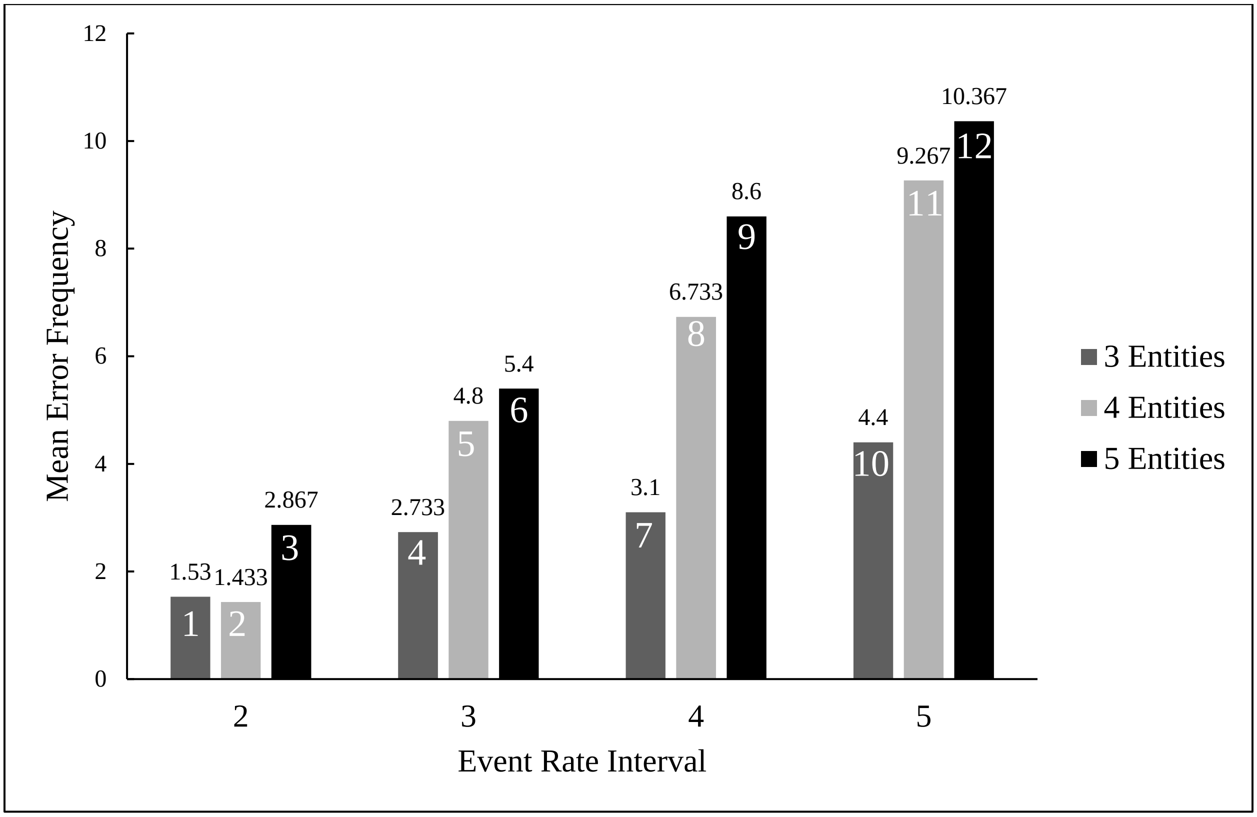


Figure G33. Mean error frequency as a function of event rate and entity number.

Table G16. Mean error frequency event rate and entity number performance comparisons matrix.

1	2	3	4	5	6	7	8	9	10	11	12
1	1		1			1					
2	2		2			2					
		3	3			3			3		
4	4	4	4			4					
				5	5				5		
				6	6		6		6		
7	7	7	7			7					
					8		8	8			
							9	9		9	9
		10		10	10				10		
								11		11	11
								12		12	12

Table G1 represents the comparison results in a format intended to provide a way to determine where equivalent accuracy performance should not be reliably expected among the different combinations of event rate and entity number. The first row of the table shows labeling which corresponds to the conditions numbered in Figure G3. The remainder of table rows indicate comparisons between the data for number indicated in the row and the remaining 11 levels. The empty cells shaded green represent significantly different comparisons from the parent number of that row ($p < .05$). All the numbered cells within a row are comparisons where the null hypothesis failed to be rejected.

The comparison presented in Table G1 gives some evidence for event rate and entity number combinations where equivalent performance should not be expected. Of particular interest in the current research are the paired comparisons to the 4 entity condition with an event frequency between 2.94 and 5.88s (labeled 3) which is indicated as condition 5 in this analysis. This condition is particularly important as it was selected as the default condition for the second experiment in this research. As shown in Table G1, the error rate for this condition was not statistically different from the error rate for conditions 6 and 10. Therefore, this analysis would indicate that the 5 entity condition with a frequency of between 2.94 and 5.88s, as well as the 3 entity condition with an event rate between 1.44 and 2.88s would have presented error rates that were not significantly different from the error rate for the selected condition.

Bibliography

Air Land Sea Application Center. (1997). *J-FIRE multiservice procedures for the joint application of firepower* (TRADOC-MCCDC-NDC-AFDC). Air Land Sea Application (ALSA).

Braze, D., Shankweiler, D., Ni, W., & Conway Palumbo L. (2002). Reader's eye movements distinguish anomalies of form and content. *Journal of Psycholinguistic Research*, *31*(1), 25-44.

Broadbent D. E., & Broadbent, M. P. (1987). From detection to identification: Response to multiple targets in rapid serial visual presentation. *Perception and Psychophysics*, *42*(2), 105-113.

Brown, H., & Prescott, R. (2006). *Applied mixed models in medicine, 2nd edition*. John Wiley & Sons, New York, NY.

Castelhano, M. S., & Muter, P. (2001). Optimizing the reading of electronic text using rapid serial visual presentation. *Behavior & Information technology*, *20*(4), 237-247.

Chapman, J. C., & Cook, S. (1923). A principle of the single variable in a speed of reading cross-out test. *Journal of Educational Research*, *8*, 389-396.

Cornsweet, T. N. (1970). *Visual perception*. Academic Press, New York, NY.

De Bruijn, O. & Spence, R. (2000). Rapid serial visual presentation: A space-time trade-off in information presentation. *Proceedings of the Working Conference on Advanced Visual Interfaces (AVI 2000 Palermo, Italy)*, New York, NY: ACM Press, 189-192.

Department of Defense (2005). *MIL-STD-2525: Department of Defense interface standard: Common warfighting symbology*. U.S. Department of Defense, Washington DC, Revised Change 1 edition, July.

Department of Defense (2010). *Technology readiness levels in the Department of Defense (DoD).* Retrieved from https://www.army.mil/e2/c/downloads/404585.pdf

Duchnicky, R. L. & Kolers, P. A. (1983). Readability of text scrolled on visual display terminals as a function of window size. *Human Factors*, *25*, 683-692

Dyson D. C., & Haselgrove M. (2001). The influence of reading speed and line length on the effectiveness of reading from a screen. *International Journal of Human-Computer Studies*, *54*, 585-612.

Federal Aviation Administration (1987). *Advisory circular 25-11 transport category airplane electronic display systems, AC 25-11*, U.S. Department of Transportation, Washington, DC.

Federal Aviation Administration (2018). Retrieved from https://www.faa.gov/nextgen/how_nextgen_works/new_technology/automation/in_depth/

Federal Aviation Administration (2009). Retrieved from https://www.esc.gov/monroneynews/archive/Vol_3/CB/09_5.asp

Forster, K. I. (1970). Visual perception of rapidly presented word sequences of varying complexity. *Perception and Psychophysics*, *8*(4) 215-221.

Frantzdale, B. (2007). Retrieved from https://www.flickr.com/photos/benfrantzdale/1056282822

Geiselman, E. E., Heft, E. L., and Miller, M. E. (2020). Development and validation of a secondary task environment for assessing visual-psychomotor tasks. *Theoretical Issues in Ergonomics Science*. doi: 10.1080/1463922X.2020.1729445

Geiselman, E. E., & Miller, M. E. (2019). The development and evaluation of a spatially-constrained text access method for information visualization. *Proceedings of IMAGE Society 2019 Conference*, 176-189. Dayton, OH.

Gawron, V. J. (2008). *Human performance, workload, and situation awareness measure handbook, 2nd edition.* CRC Press, New York, NY.

Intraub, H. (1980). Presentation rate and the representation of briefly glimpsed pictures in memory. *Journal of Experimental Psychology: Human Learning and Memory, 6*(1), 1-12.

Intraub, H. (1981). Rapid Conceptual identification of sequentially presented pictures. *Journal of Experimental Psychology: Human Learning and Memory*, *7*(3), 604-610.

Just M. A., & Carpenter P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*(4), 329-354.

Kahneman, D. (1973). *Attention and effort*. Prentice-Hall, Englewood Cliffs, NJ.

Kantowitz, B. H., & Sorkin, R. D. (1983). *Human factors*. John Wiley, New York, NY.

Keppel, G. (1992). *Design and Analysis: A Researcher's Handbook.* Prentice-Hall, Englewood Cliffs, NY.

Kruk, R. S., & Muter P. (1984), Reading of continuous text on video screens. *Human Factors*, *26*(3), 339-345.

Le Grand, Y. (1975). Measurement of the visual stimulus. In Carterette, E. C. & Friedman, M. P (Eds.), *Handbook of perception volume V: Seeing*. Academic Press, New York, NY.

Lindsay, P. H., & Norman, D. A. (1972). *Human Information Processing*, Academic Press, New York, NY.

Matin, E. & Boff, K. R. (1988) Information transfer rate with serial and simultaneous visual display formats. *Human Factors*, *30*(2), 171-181.

Morrison, R. E., & Rayner, K. (1981). Saccade size in reading depends upon character spaces and not visual angle. *Perception & Psychophysics, 30*, 395-396.

Muter, P. (1996). Interface design and optimization of reading of continuous text. In *Cognitive aspects of electronic text processing* (pp. 161-180). Ablex, Norwood, NJ.

Nieuwenstein, M. R., & Potter M. C. (2006). Temporal limits of selection and memory encoding: a comparison of whole versus partial report in rapid serial visual presentation. *Psychological Science, 17*(6), 471-475.

Oksama, L., & Hyönä, J. (2004). Is multiple object tracking carried out automatically by an early vision mechanism independent of higher-order cognition? An individual difference approach. *Visual cognition*, *11*, 631–671.

Öquist, G. (2006). *Evaluating readability on mobile devices*. Acta Universitatis Upsaliensis. Studia Lingustica Upsaliensia, Uppsala. ISBN 91-554-6745-8.

Osgood, S. S., Boff, K. R., & Donovan, R. S. (1988). Rapid communication display technology efficiency in a multi-task environment. *Proceedings of the Human Factors Society 32nd Annual Meeting*, 1395-1399.

Paterson, D. G., & Tinker, M. A. (1929). Studies of typographical factors influencing speed of reading, II. Size of type. *Journal of Applied Psychology, 13*(2), 120–130.

Paterson, D. G., & Tinker, M. A. (1940). Influence of line width on eye movements, *Journal of Experimental Psychology*, *27*, 572-577.

Payne, D.G., & Lang, V.A. (1991). Visual monitoring with spatially versus temporally distributed displays. *Human Factors 33*(4), 443-458.

Piolat, A., Roussey, J., & Thunin, O. (1997). Effects of screen presentation on text reading and revising. *The International Journal on Human-Computer Studies, 47*, 565-589.

Potter, M. C. (1984). Rapid serial visual presentation (RSVP) a method for studying language processing. In D. E. Kieras and M. A. Just (Eds.), *New Methods in Reading Comprehension Research* (pp. 91-118). Erlbaum, Hillsdale, NJ,

Pylyshyn, & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, *3*, 179-197.

Rahman, T. and Muter, P. (1999). Designing an interface to optimize reading with small display windows. *Human Factors*, *41*(1), 106-117.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 374-422.

Richardson, T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, *6*(2), 135-147

Robeck, M. C. & Wallace, R. R. (1990). *The psychology of reading: an interdisciplinary approach, 2nd edition*. Lawrence Erlbaum Associates. Hillsdale, NJ.

Roscoe, A. H. & Ellis G. A. (1990). *A subjective rating scale of assessing pilot workload in flight: a decade of practical use*, *a technical report.* Procurement Executive, Ministry of Defence, Farnborough, Hampshire.

Roth, V. (2008). Diagonal reading, or don't lose your line. *Proceedings of the 16th annual ACM international conference on multimedia*, ACM, New York, NY.

Spence, R. (2002). Rapid, serial and visual: a presentation technique with potential. *Information Visualization*, *1*, 13-19.

Tinker, M. A., & Paterson, D. G. (1929). Studies of typographical factors influencing speed of reading. III. Length of line. *Journal of Applied Psychology, 13*(3), 205–219.

U.S. Air Force (2011). Retrieved from https://media.defense.gov/2011/Dec/08/2000193354/-1/-1/0/111205-F-ZZ987-004.JPG.

Vandierendonck, A. (2018). Further tests of the utility of integrated speed-accuracy measures in task switching. *Journal of Cognition*, *1*(1), 8.

Wickens C. D., Lee, J. D., Liu Y. & Gordon Becker S. E. (2004). *An Introduction to human factors engineering, 2nd ed.*, Pearson Prentice Hill. Upper Saddle River, NJ.

Wickens, C. D. (2008). Multiple resources and mental workload. *Human factors*, *50*(3), 449-455.

Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2013). *Engineering psychology and human performance, 4th ed.*, Pearson. Boston, MA.

Vita

Name: Eric E. Geiselman

Duty (PD) Title: Senior Engineering Research Psychologist

Area of Expertise: Visual Display Design

Education:

- M.A. Experimental Psychology/Human Factors, University of Dayton, 1991.
- B.S. Psychology, University of Dayton, 1988.

Specialized Training:

- Doctoral Candidate, Systems Engineering, Air Force Institute of Technology, expected graduation: Mar 2020.
- Commercial Pilot: Single and Multi-Engine Land Airplane – Instrument, CL-65 Turbojet Type Rating.
- 7500 hours total flight time logged.
- Certified Acquisition Professional, Level III, Test and Evaluation.

Significant Accomplishments: **Total Number of Peer-Reviewed: Publications** 50

- Designer of the Airdrop Guidance Display (ADG) in support of precision non-guided airdrop tasks.
- Designer of the "Theta" transparent ownship attitude display (X-31 and JSF flight tests).
- Designer of the "reflected" line-of-sight oriented locator line display for tactical target tracking (JSF flight test).
- Designer of the Non Distributed Flight Reference Display (NDFR).
- Developed and tested symbology color-coding strategies for the Helmet-Mounted Sight Plus program.

Selected Peer-Reviewed Publications:

- Geiselman, E.E, Heft, E., & Miller, M.E. (2020). Development and validation of a secondary task environment for assessing visual-psychomotor tasks. *Theoretical Issues in Ergonomics Science*, Taylor & Francis.
- Geiselman, E. E., Quill, L. L., Cox, N. J., & Dubois, J. A. (2016). Airdrop guidance display format for a precision airdrop application on an auxiliary display equipped aircraft. *International Journal of Aviation Psychology, 25*(3), 141-156.

- Geiselman, E. E., Johnson, C. M., Buck, D. R. & Patrick, T. (2013) Flight deck automation: a call for context-aware logic to improve safety. *Ergonomics in Design: The Quarterly of Human Factors Applications*, *21*(4), 13-18.
- Geiselman, E. E., Johnson, C. M., & Buck, D. R. (2013) Flight deck automation: Invaluable collaborator or insidious enabler, *Ergonomics in Design: The Quarterly of Human Factors Applications, 21*(3), 22-26.
- Geiselman, E.E. and Osgood, R.K. (1995). Helmet-mounted display attitude symbology: an evaluation of compression ratio. *International Journal of Industrial Ergonomics, 15*, 111-121.

Significant Awards/Honors:

- AFMC-level selectee: 68th Annual Arthur S. Flemming Award: Applied Science and Engineering, 2016
- Fitts Human Engineering Award, 2015
- Air Force Research Laboratory Human Effectiveness Directorate Civilian of the Quarter: March, 2013

Memberships with Professional/Technical Societies:

- Human Factors and Ergonomics Society (HFES)
- Southern Ohio Chapter of the Human Factors and Ergonomics Society (SOHFES)
- Department of Defense Human Factors Engineering Technical Advisory Group (DOD HFE/TAG)

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
26-03-2020	Doctoral Dissertation	April 2017 – March 2020

4. TITLE AND SUBTITLE
Evaluation of a Text-Based Information Portrayal and Techniques for Constrained Area Presentation:
A Comparison of Human Performance

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Geiselman, Eric, E., GS-14, 711th HPW/RHCSR, USAF

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Institute of Technology
Graduate School of Engineerin(
Management (AFIT/EN)
2950 Hobson Way
WAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT
AFIT-ENV-DS-20-M-205

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AFRL 711 HPW/RHCSR
Section Chief Brown,
BLDG 441
WPAFB, OH 45433

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL 711HPW/RHCSR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
DISTRIBUTION STATEMENT A. APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

13. SUPPLEMENTARY NOTES

14. ABSTRACT
To maximize visualization effectiveness, graphical data are commonly augmented with alphanumeric (i.e., text) symbols to provide detailed information and define specific values. This text is often provided through a pop-up dialog box which contains all of the alphanumeric data pertaining to an object simultaneously. However, a human performance cost can be associated with the current portrayal and access techniques as the resulting portrayal can occlude critical information in the visualization and degrade operator performance. This research included the development and evaluation of three alternate spatially-constrained text portrayal techniques. These techniques and their associated access interface were designed to reduce occlusion while providing rapid access to desired alphanumeric data. The techniques were comparatively evaluated using a dual-task human performance paradigm. Among the performance measures were accuracy, response time, and subjective feedback. The basis for development of the spatially-constrained text access techniques is discussed as well as their implementation affordances and limitations. The spatially-constrained text portrayal and access user interface concepts resulted in mixed accuracy and response time performance compared with the more conventional method. Specific design features promoted a 3X reduction in data access time versus unaided spatially-constrained text portrayal. Overall, it was shown that equivalent performance was obtained among the variants while the potential for occlusion was reduced during use of the novel designs.

15. SUBJECT TERMS
secondary task; visual information processing; standardised task

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Dr. Michael E. Miller,
U	U	U	UU	295	1. **19b. TELEPHONE NUMBER** *(include area code)* (937) 255-3636 x4651 Michael.Miller@afit.edu

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18