EVALUATION OF DIGITAL CHECKLISTS FOR COMMAND AND CONTROL OPERATIONS

Christopher K. McClernon[1], Victor S. Finomore[2], Terence S. Andre[3], Forrest S. Jeffery[1], & Oliver N. Myers[1]

United States Air Force Academy, CO[1]
Air Force Research Laboratory, Wright-Patterson AFB, OH [2]
TiER1 Performance Solutions, Covington, KY [3]

Checklists are a critical component of most any military operation and both the construction of and presentation of checklists can affect the performance and efficiency of these operations. The purpose of this research project is to compare methods for displaying and executing checklists in a command and control operation to increase both performance and efficiency. The NORAD/NORTHCOM Command Center (N2C2) uses a paper checklist system to facilitate responding to any number of disasters, to include air catastrophes. This project investigated the potential effectiveness of a digital system that could take the place of the paper system that is currently being used. A between groups experimental design was used to analyze the relative effectiveness of each method. Each group of subjects was introduced to a timed task on different checklist systems and asked to complete an Air Force Academy Command Center checklist as accurately and as quickly as possible. Performance and subjective assessments of each system were analyzed and compared. The data showed that a linear digital checklist takes a longer amount of time than both a paper checklist and hierarchical digital checklist. In addition, the subjective data showed that the hierarchical system and paper system were easier to use and navigate than the linear system.

INTRODUCTION

The NORAD/NORTHCOM Command Center (N2C2) in Colorado Springs, Colorado is responsible for nation-wide military responses to national disasters. For the past 30 years, NORAD has been using paper-based checklists to respond to national disaster emergencies. The checklist binder consists of hundreds of pages divided into 7 sections dealing with different types of national emergencies (maritime, land, air, etc). Operators search for the appropriate checklist(s) and follow instructions based on the type of emergency situation that occurs. N2C2 expressed an interest in integrating a new digital system in place of the paper checklist if the new design will benefit the operational crews. While numerous studies explore electronic checklist use in aviation (Mosier, Palmer, & Degani, 1992; Palmer & Degani, 1991), medical (Hales & Pronovost, 2006; Ko, Turner, & Finnigan, 2011), engine room (Kluj, 1999), and space (Marmolejo, 1996) applications, there is a dearth of literature on checklists in a command and control setting. Empirical data is needed to show what type of system will create the easiest transition through training while allowing command center operators to complete the checklist in a shorter amount of time with a limited amount of errors.

Hewett et al. (1992) defines human-computer interaction (HCI) as a "discipline concerned with the design, evaluation and implementation of interactive computing systems for human use as well as the study of the major phenomena surrounding them" (p. 5). They also identify key characteristics of HCI such as the structure of communication between human and machine as well as the importance of thorough analysis of the user tasks when designing an effective interface. These characteristics are critically important when designing a checklist of any kind, especially for a mission as important as N2C2. NORAD's current operations consist of different teams working together to complete each checklist appropriate for the emergency. In order for the system to contribute to efficient operations, it must enhance the interaction between the human and the system. A newly implemented system cannot decrease efficiency of operations.

Wickens, Gordon, and Liu (2004) discuss software usability and state that the criteria is determined by the system's learnability, efficiency, memorability, error rate, and user satisfaction. With a newly integrated system, users will start as novices but will be trained and can become expert users. Extensive training on a new system will ensure a lower error rate but this can only be accomplished if the learnability, memorability, and user satisfaction of the new system are acceptable. In the case of the N2C2, users of the paper system are trained and accustomed to that system; if a new system is introduced, the users must be trained on the system in order to get them to the level of proficiency that was previously achieved with the paper system. In order to replace the current system, the new system must have a high learnability, memorability, and low error rate while decreasing the amount of time that it takes to complete tasks.

According to Shneiderman & Plaisant (2009), previous HCI research shows that the quality of the user interface has a direct impact on a user's efficiency and satisfaction. The practical aim of research in HCI is to improve interface design (Carroll, 1987). Designers must make computer systems easier for humans to learn, use, and understand. A faulty HCI design traps the user into "unintended and mystifying circumstances." (Gerlach and Kuo, 1991). The designers of any new system that may be implemented at NORAD must take into consideration that the system must be easily learnable. Transitioning to a system that takes more mental workload

than the previous system will negatively impact user satisfaction and possibly performance on the system as well.

Background

To fully understand the N2C2 system and the current checklist method employed, the authors traveled to NORAD and met with senior leaders in the N2C2. During the visit, data was collected in the form of interviews, observational data, and usability analysis of existing systems. Results determined that the existing paper checklist is effective, but can be drastically improved. Operators demonstrated their cumbersome method for navigating to numerous checklist locations concurrently while directing numerous response organizations.

An investigation into the N2C2 concept of operations (CONOPs) also revealed insight into why a redesign of the system's checklists is advantageous. One of the N2C2 primary objectives is to shorten response times and improve collaboration for routine and emergency operations. Other objectives mentioned include sharing of information, flexibility, electronic log data, reduce operator workload, and continuity between various centers. All of these objectives support the introduction of a modern electronic checklist to replace the legacy paper system. (NORAD, 2011)

The authors then investigated numerous methods for presenting electronic checklists. Pilot studies gathered performance and usability data from various platforms, and the two formats, linear and hierarchical, were selected for testing in this project. Linear checklists are the most common form of electronic checklists. They are digital documents that may be manipulated and navigated through a common menu structure. Typically a side bar allows navigation at various speeds, making them advantageous for navigation through large documents. However, as the desired location is reached, tedious scrolling of the screen can appear laborious at times. Hierarchical checklists are more interactive, in that hyperlinks may transport readers quickly to relevant sections. A very simple example is a clickable table of contents. These two formats in addition to a paper checklist were tested in this study.

Purpose

The aim of this study was to evaluate the effectiveness and usability of three systems. The N2C2 will be able to use the data found and look into more studies to be done in the area while implementing some of these findings into their operations. These results and the development of a new N2C2 checklist system could ultimately benefit our national security.

METHODS AND MATERIALS

Participants

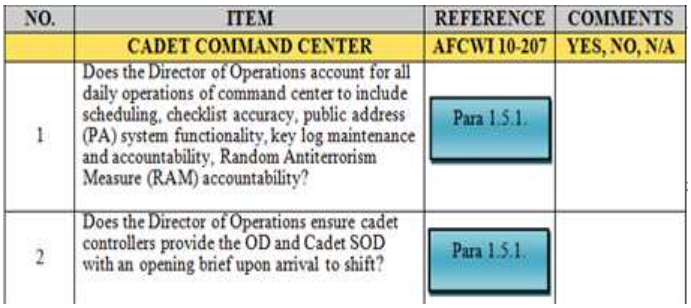
Participants (N=21) were recruited from the Air Force Academy participant pool (age 18-23). The participants had a similar male to female proportion (F=4, M=17) as the Academy and the Air Force populations. Participants were randomly assigned to a linear, hierarchical, or paper checklist group. All of the cadet participants are future Air Force officers, and are therefore generally representative of Air Force members. None of the participants had experience with the cadet command center checklist used in this study, nor were they familiar with command center operations.

Equipment

A 16-item command center checklist was utilized for this study. Each of the 16 items required users to lookup pertinent regulations regarding Academy crisis response items, and then users were required to either check the item as complete or fill in required information (e.g., Was the area secure?). The content of each checklist was identical, and the only difference was the presentation method:

Paper – The printed paper checklist was representative of a tradition, hard-copy paper format. Navigation was accomplished by turning tabbed pages, and checklist items were accomplished using a pencil. This presentation method is representative of the checklist currently utilized by the N2C2.

Linear – A digital form of the paper checklist (.pdf) was presented on a laptop computer, and a computer mouse scroll wheel was used to navigate the checklist (reference Figure 1). Mouse clicks identified accomplished checklist items, and the computer keyboard was used to enter required information.



NO.	ITEM	REFERENCE	COMMENTS
	CADET COMMAND CENTER	AFCWI 10-207	YES, NO, N/A
1	Does the Director of Operations account for all daily operations of command center to include scheduling, checklist accuracy, public address (PA) system functionality, key log maintenance and accountability, Random Antiterrorism Measure (RAM) accountability?	Para 1.5.1.	
2	Does the Director of Operations ensure cadet controllers provide the OD and Cadet SOD with an opening brief upon arrival to shift?	Para 1.5.1.	

Figure 1. Linear Checklist used for data collection

Hierarchical – A digital checklist was completed, but navigation was accomplished with hyperlinks between items. The appearance of the hierarchical checklist was the same as the linear checklist (see Figure 1) except the "REFERENCE" column contained hyperlinked boxes to relevant sections.

A stopwatch was used to record the time to answer each checklist item as well as overall checklist completion time. In addition, the IBM Usability Questionnaire instrument (Lewis, 1995) was used to examine qualitative assessments of the instruments.

Manipulations

A between groups experimental design was used to analyze the relative efficiency of each checklist system. Each group of participants were randomly assigned to a condition and introduced to a timed task on one of the different systems

(linear, hierarchical, or paper) and asked to complete an Air Force Academy Command Center checklist as accurately and as quickly as possible. Performance and subjective assessments were collected and compared.

Procedures

Cadets were first introduced to the system that they were randomly assigned to. They were then given a short training overview of the checklist they were assigned to and how the checklist works. The participants were then asked to complete the assigned checklist as quickly and as accurately as possible. Each checklist consisted of the same 16 questions. The questions were relevant to security of the Academy's cadet area and required participants look up and record answers from relevant regulations. This format was used to simulate the tasks and actions required of an N2C2 commander following the execution of a checklist item.

As each person completed the items on the checklist, the amount of time it took them to move on to each item on the checklist was marked so that the time between items could be measured and compared. Measuring the time each item took individually allowed for questions of different lengths to be compared between the systems since longer questions most likely take longer to answer. The total time that it took each participant to complete the checklist was also recorded in addition to the number of errors that each participant made while answering checklist items. After each participant completed the timed checklist, they were asked to fill out the IBM Usability Questionnaire about their satisfaction of the system they used.

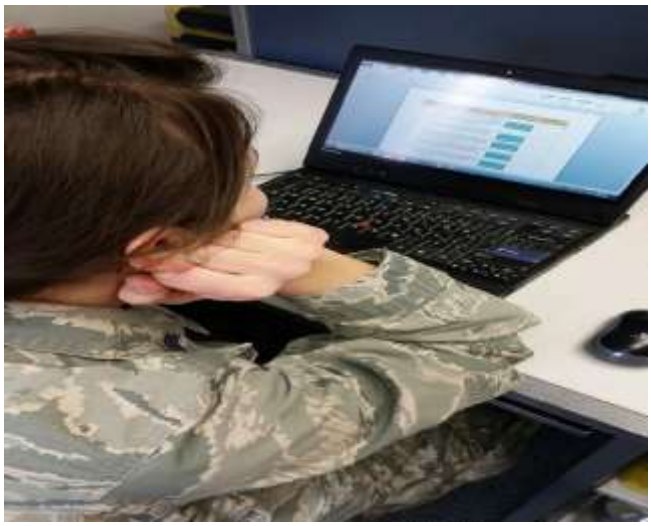


Figure 2. A cadet participant completing the linear checklist

Measures

In this experiment, task performance was measured by how long each participant took to answer the entirety of the checklist as well as how long each of the 16 items took to be answered individually. Another measure was the accuracy of each participant's checklist answers. The final measure was the subjective data gathered by each participant filling out the IBM Usability Survey.

RESULTS

Performance

Mean time spend on each item on the checklist is displayed in Figure 3 for all three experimental conditions. As seen in Figure 3, participants in the Linear condition took longer to complete the items then participants in the Hierarchical and Paper conditions. A one-way ANOVA identified a statistically significant difference across the checklist, F (2, 17) = 9.03. p < .05. Post Hoc Tukey test revealed that participants assigned to the hierarchical and paper checklists accomplished the task significantly faster than participants assigned to the linear checklist and that there was no difference in completion time between the Hierarchical and Paper checklists.

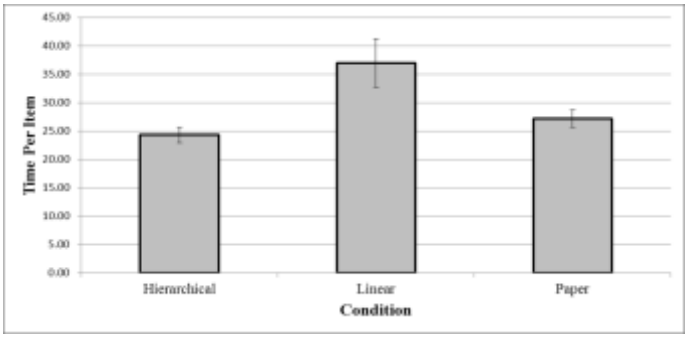


Figure 3. The linear checklist took longer to complete compared to the hierarchical and paper checklists.

When analyzing the accuracy of checklist responses, no significant differences were found between any of the three checklist methods (p>0.05).

Qualitative data collected by the IBM Usability Survey after the checklists were completed showcases that that participants preferred the usability of the hierarchical and paper checklist (see Figure 3). An ANOVA confirmed a statistically significant difference between the participants' usability scores across conditions, F (2, 17) = 6.712, p < 0.05. Post Hoc Tukey, found that participants rated the linear conditions as the least useable condition and that there was no difference between Hierarchical and Paper checklists.

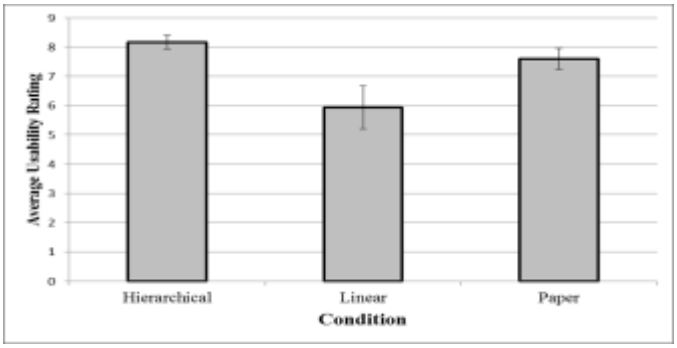


Figure 4. Users preferred the hierarchical and paper checklist formats compared to the linear checklist

DISCUSSION

This study determined that the hierarchical and paper checklists are more efficient compared to a linear checklist. However, no analyses revealed any difference in accuracy between the three checklist methods. This is especially important in an N2C2 context--these operators currently use a checklist comparable to the paper checklist in this study. By varying the presentation method of the checklist, these results indicate that changes in accuracy may not be observed, and they may perform the checklist as quickly when changing from a paper checklist to a hierarchical checklist.

The participants also preferred the usability of the hierarchical and paper checklists compared to the linear checklist. These results are also in light of the variance expected in a between subject design. Participants appear to not like the usability of the linear checklist that had to be scrolled through and which resulted in this group taking the longest time to complete the checklist. These results are not surprising given the rather laborious navigation method attributed to a linear method of checklist presentation.

While the linear checklist fared worse in most regards, the similar results between the hierarchical and paper checklists beg the question: Why should the N2C2 (or other crisis response center) change from a paper to hierarchical checklist if they are equal in speed, accuracy, and usability? In interviews with the N2C2 operators, the paper navigation is very cumbersome. Operators must hold multiple places within hundreds of checklist pages using many of their fingers. We predict that our relatively short and simple Academy checklist did not require these inconveniences and, therefore, gave the paper checklist an advantage in our study when compared to the N2C2. The hierarchical checklist may also be superior in terms of customization and revisions. A current disadvantage of the N2C2 paper checklist is that it requires printing, and then locating the appropriate page(s) in the numerous control stations throughout the N2C2 network when revisions are made. In addition, there is a need for operators to carry the checklist around the control center with them as they direct various military assets. The current checklist is not very mobile, and a digital checklist could greatly enhance portability of the checklist.

Another difference between the N2C2 and Academy checklists is the urgent nature of each checklist. Cadet participants in this study were in a low threat environment and not responding to a national emergency or disaster. In contrast, the N2C2 operators are responding to real world crisis events that may potentially threaten our national security. Because of this, there may be great psychological disparity between the two groups of operators. Further testing of these various checklist configurations are being explored in various operational contexts.

Some research demonstrating a negative correlation between age and comfort with computers (Czaja & Sharit, 1998) may suggest that younger users prefer a digital or technology-based checklist solution when compared to older users. N2C2 commanders that use the crisis response checklist are typically senior military officers with many years of military experience, while the cadet participants used in this study are college-aged students (18-23 years old). The fact that the paper checklist faired so well in this study with younger users is especially unexpected, and may further support that the checklist used in this study was over-simplified. In addition, computer comfort may be highly modifiable with computer usage of all age groups. Future studies using a larger range in participant ages are currently being explored.

The current study also did not test the learnability of the checklist presentation methods. Future work is underway to determine if any of the checklist presentation methods is easier to learn for novice users. Likewise, the current work did not test the speed, accuracy, and usability of a well-practiced checklist (i.e., expert users). We plan to test the performance of checklists following a rigorous training program on each checklist.

Technological advances allow innovative solutions to checklists in numerous contexts. For example, in the aviation industry, numerous military and civil pilots are resorting to electronic checklists to accomplish the myriad of complex tasks required in this context. Likewise, numerous healthcare professionals are resorting to electronic checklists for the benefits highlighted in this study (Hales & Pronovost, 2006; Ko, Turner, & Finnigan, 2011). However, very little work is examining electronic checklists in a command and control setting. Given the incredibly important mission of N2C2 and the impact on our national security, an alternative, technologically advanced means for managing commander's checklists may be necessary.

CONCLUSION

The goal of this study was to empirically determine the best way to present checklists out of three different designs so that we could provide warfighters with a system that can replace the N2C2 paper checklist with an electronic system. Our data shows that there is a difference in performance with respect to time but not accuracy when comparing linear, paper, and hierarchical checklists. These results highlight the possibility for NORAD/NORTHCOM to replace their paper checklist while not losing efficiency and accuracy in their operational response to national disasters. The quality of a user interface has a direct impact on efficiency and satisfaction with the system. Since the paper checklist has been used for the past 30 years, operators in the N2C2 are likely more comfortable with a paper system and may not initially be open to the idea of a hierarchical design of a digital system. When presented with this data, they might begin to feel differently. Through experience and practice on a new system, a hierarchical system may prove to be an equal or more efficient system in regards to the NORAD/NORTHCOM mission of national disaster response. These findings can now be shown to others and replicated at higher levels. Since not much research has been done for paper to digital systems, this data can contribute to that field and provide insight into transitioning between the two formats.

THANK YOU TO OUR SPONSORS

The authors would like to thank the USAF Academy's Warfighter Effectiveness Research Center, the Air Force

Research Laboratory's Human Effectiveness Directorate, TiER1 Performance Solutions, and the Air Force Office of Scientific Research for their continued support. The work presented in this paper fulfilled the research requirements for the Human Factors degree in the USAF Academy's department of Behavioral Sciences and Leadership.

REFERENCES

Carroll, J. M. (1987). *Interfacing thought: Cognitive aspects of human-computer interaction*. Cambridge, MA: MIT.

Czaja, S. J., & Sharit, J. (1998). Age differences in attitudes toward computers. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *53*(5), 329-340.

Gerlach, J. H. and Kuo, F. (1991). Understanding human-computer interaction for information systems design. MIS Quarterly. 15 (4), 527-549.

Hales, B. M., & Pronovost, P. J. (2006). The checklist—a tool for error management and performance improvement. *Journal of critical care*, *21*(3), 231-235.

Hewett, T. T., Baecker, R., Card, S., Carey, T., Gasen, J., Mantei, M., & Verplank, W. (1992). *ACM SIGCHI curricula for human-computer interaction*. ACM.

Kluj, S. (1999). Use of checklists in engine room simulator. In *Proceedings of 4th International Conference on Engine Room Simulators, Vallejo*.

Ko, H. C., Turner, T. J., & Finnigan, M. A. (2011). Systematic review of safety checklists for use by medical care teams in acute hospital settings-limited evidence of effectiveness. *BMC health services research*, *11*(1), 1.

Lewis, J. R. (1995). IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, *7*(1), 57-78.

Marmolejo, J. A. (1996). *An Electronic Cuff Checklist Information Display for Extravehicular Activity* (No. 961527). SAE Technical Paper.

Mosier, K. L., Palmer, E. A., & Degani, A. (1992, October). Electronic checklists: Implications for decision making. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 36, No. 1, pp. 7-11). SAGE Publications.

NORAD (2011). *Battle staff standard operating procedures* (No. 1-01). NORAD & USNORTHCOM Publication Series.

Palmer, E., & Degani, A. (1991, April). Electronic checklists: Evaluation of two levels of automation. In *Proceedings of the Sixth Symposium on Aviation Psychology* (pp. 178-183).

Shneiderman, B., & Plaisant, C. (2009). Human factors of interactive software. *Designing the user interface strategies for effective human-computer interaction* (5th ed., pp. 29-30). MD: Prentice Hall.

Wickens, C. D., Gordon, S. E., & Liu, Y. (2004). Human-computer interaction. *An introduction to human factors engineering* (pp. 453-456). New York City, NY: Pearson Prentice Hall.