AUTOMATION IN THE HOME: THE DEVELOPMENT OF AN APPROPRIATE SYSTEM REPRESENTATION AND ITS EFFECTS ON RELIANCE

Julian Sanchez, Gina Calcaterra, & Quan Q. Tran
Georgia Institute of Technology
Atlanta, Georgia

To achieve an appropriate level of reliance on an automated system, the operator must have an accurate system representation such that he/she is aware of the capabilities and limitations of the system. The appropriate use of an automated system can lead to optimal performance by the human-machine dyad. This study investigates the relationship between an accurate system representation and behaviors associated with human-automation interaction (e.g., reliance). Furthermore, age-related effects are also included in the investigation. A cooking memory aid (Cook's Collage) is used as an automated aid that keeps track of the ingredients used in a specific recipe. Participants are asked to interact with the automated device for 5 sessions (each on different days). Tasks are structured to simulate those in a real kitchen. Preliminary results suggest that that there is a different pattern of interaction with the aid as a function of age. Older adults tend to rely on the aid for real-time feedback while younger adults use the aid as a verification tool that they have executed the recipe as desired.

INTRODUCTION

Automated systems are always designed to operate within a specific set of parameters, and if used appropriately within those parameters the automation usually behaves in a reliable fashion. An issue in the study of human-automation interaction is the extent to which the automation is relied on by the human. Ideally, the human should appropriately use the automation, which means they should only rely on it in situations that the automation is designed to handle and will provide useful support. If the human uses the automation in situations in which it may prove to be unreliable, then it is being misused. Conversely, if they are not using the automation in circumstances in which it is reliable then it is being underused or disused (Parasuraman & Riley, 1997). To achieve appropriate use, perceived reliability has to match the actual reliability of the system; more specifically, the human has to understand the capabilities and limitations of the automation and use it accordingly.

Matching Perceived and Actual Reliability

Automation failures may be caused by the way in which the automation is modeled and designed. In specific conditions, the aid or support provided by the automation may not meet the needs of the system. Changes in operator priorities, imprecision by sensors caused by unforeseen conditions, or changes in the operational environment are some of the circumstances that may lead to situations where the algorithms employed by the automation are inadequate (Cohen, Parasuraman, & Freeman, 1999; Lee & See, 2004). This component of reliability, referred to as capabilities of the automation, denotes the probability that the automation is working properly or "behaving" such that it positively contributes to the accomplishment of a goal or a task successfully. Unreliable automation, as defined by a shortage in the capabilities of the system, has been referred to as "imperfect automation" (Wickens, Lee, Liu, & Gordon-Becker, 2004). Imperfect automation consistently and accurately performs all of the functions it was designed to perform but it may not always perform all the functions according to the task criterion. The task criterion serves as the standard to determine if the advice or support provided by the automation is accurate.

Automation and Older Adults

The issue of how older adults use and rely on automated systems has not received much attention. A few studies have investigated the issues that might be pertinent to how older adults adapt and use automated support. For example, Sanchez, Fisk and Rogers (2004) found that older adults were as likely to rely on automated status indicators as younger adults in a multi-task environment. They also found that older adults were more perceptive to the drops in the reliability of the automation. Dingus et al. (1997) found that the increased presence of false alarms in a driving scenario did not affect the driving behavior of older adults.

In a study of visual detection in a luggage screening task, McCarley, Wiegmann, Wickens, and Kramer (2003) found that younger adults benefited from an automated aid by increasing their sensitivity relative to the non-automated condition. However, older adults' detection sensitivity did not increase with the presence of the automated aid. Interestingly, the perceived reliability estimates did not differ as a function of age. This study showed that even when both younger and older adults' perceived reliability estimates were similar there were still differences in how the automation was used. Some have suggested that the acceptance and likelihood of use of new technologies decreases with age (Kantowitz, Becker & Barlow, 1993).

Automation in the Home

As the use and implementation of automated agents becomes more prominent across a variety of operational environments, many of the issues that have been observed in human-automation interaction in contexts such as aviation and surface transportation may present themselves in environments such as the home. These issues include ones such as reliance and trust in the automation. Currently there are a number of research initiatives with the aim of investigating solutions to such issues within the home. An example of one of these research efforts is the Aware Home Research Initiative at the Georgia Institute of Technology (Mynatt, Melenhorst, Fisk, & Rogers, 2004; www.awarehome.gatech.edu).

The Aware Home is an "intelligent" house designed to assist with a variety of tasks including support for Activities of Daily Living (ADLs) such as cooking and recognition and detection of emergency situations such as fire detection (Sanchez, Ezer, Rogers & Fisk, in press). One example of a system being used to study human-automation interaction in the home is the Cook's Collage (Figure 1). This system is designed to serve as a memory aid in multi-task environments such as the kitchen. The system records hand movements linked with specific ingredients and provides memory support by way of a visual record of what has been added during the preparation of a recipe. This visual record is meant to assist a person in recovering from interruptions as well as reducing workload.

THE CURRENT STUDY

The purpose of the study is to investigate if there is a relationship between participants' knowledge about the capabilities of the system and the development of attitudes and behaviors towards the system (e.g., trust, perceived reliability, and reliance). Participants are given the opportunity to develop strategies to interact with the automated device while interacting with it under a variety of conditions that resemble those that one may encounter in a real kitchen. If participants are able to adapt their strategies to the capabilities and limitations of the automated cooking aid, their performance in cooking may benefit, especially under high workload conditions (e.g., something on the stove begins to burn). It is worth noting that the reliability of the cooking aid (which is not perfect) is not systematically controlled. Like many systems in applied contexts, its reliability depends on how it is used. Therefore, in this study, participants are allowed to freely interact with the system (within the constraints of trying to accomplish a specific end-goal) and qualitative assessments are made by way of interviews to determine

Figure 1. Experimental environment

Participants

A total of 24 participants will participate in the study. Half of the participants will consist of younger adults (age 20 – 45) and the other half of older adults (60 – 80). All participants will be compensated for their time ($5/hour, with a bonus that brings total payment to $60 if all 5 sessions are completed). Currently, data is still being collected.

Apparatus

The Cook's Collage consists of two cameras that continuously record hand movements (one camera for each hand). The video is fed to an experimenter (who sits in another room) and who counts the number of movements that are associated with mixing ingredients. The experimenter relays this information in real-time back to the display that the participants interact with. It is worth noting that participants are not told that there is a human component in the system. The Cook's Collage has the following limitations:

- The accuracy of the system in keeping a true count of the number of additions per ingredient is directly affected by the speed of the participants' hand movements. The faster the participant adds repetitions of an ingredient, the more likely the count kept by the system will be inaccurate.

- The system has a temporal delay of approximately 15 seconds between the time when an ingredient is added and the time it gets counted and appears on the display.

Tasks

Cooking task. Two different recipes are used in the experiment (punch and cinnamon-sugar cookies). Both recipes require approximately the same number of steps and ingredients. The recipes are each posted on one of the kitchen cabinet doors, where participants can easily look at them when they need to. Participants are told to follow the recipe as close to specifications as possible and to carry out the recipe at a cooking pace that is comfortable for them.

Stove monitoring task. A stove monitoring simulation is used as a multi-tasking scenario. Participants are told to monitor a computer touch screen on which they are cooking three dishes. The three burners increase and decrease in temperature at different rates. The goal is to keep the temperature of the food in a neutral cooking zone. Participants are asked to prevent the items from either burning or cooling on the stove by activating a High/Low switch. They are instructed to press the "High" button if their food begins to fall into the "cooling" range and press the "Low" button if the food falls into the "burning" range. At the conclusion of the task, participants are given feedback in terms of the percentage of time their food was either burning or cooling during the given task. The purpose of this task is to introduce an active monitoring component to the experiment. This will help us evaluate the reliance on the Cook's Collage under conditions of higher workload.

Interruption task. An alarm system simulation is used to interrupt the cook while they are making a recipe. Participants are told that a friend is installing a new security system in their home, and they should help their friend out by disabling the system if the alarm happens to go off while cooking. Prior to the start of the task, the participant is told to choose a 5-digit pin number that they will remember if they should need to use it. When the alarm goes off during the task, the participant walks from the kitchen into the adjacent living room where they disable the alarm by entering their 5-digit pin number on a touch screen mounted on the wall. When an interruption does occur, the participants have a certain time window during which they should disable the alarm. Alarm interruptions occur randomly throughout the task, with a minimum of one interruption every 90 seconds. The purpose of this task is to investigate reliance on the Cook's Collage when participants have to "disengage" from the cooking task for a longer period of time.

Design

During each session, all participants perform the tasks in the following order of trials (2 and 3 are counterbalanced):

1. Cooking task (punch)
2. Cooking task (punch) with stove monitoring task
3. Cooking task (punch) with stove monitoring task and interruption task
4. Cooking task (cinnamon-sugar cookies) with stove monitoring task and interruption task

Measures

After every trial, subjective measures of trust and workload (NASA-TLX questionnaire) are gathered. The trust questions are specific to the Cook's Collage. The workload questions are used to assess how demanding each particular task is for the participant. Furthermore, participants are asked open ended questions regarding the strategies they employed to interact with the system. These questions are meant to inquire about the knowledge they have acquired that relates to the capabilities and limitations of the system. The accuracy with which the recipes are carried out is also measured.

Reaction times are collected for the interruption task. For the stove monitoring task, the percentage of time the temperature of the food was out of range (burning or cooling) is measured. Participants are asked open ended questions regarding the strategies they employed for monitoring the stove and their perceived performance on the task (how often their food was burning or cooling). Again, the objective of this analysis is to determine if there is a relationship between overall performance in the task and the human's representation of the cooking aid. It is expected that those who figure out the system's limitations will be able to adopt strategies that help them utilize it effectively.

Procedure

Participants are brought in for five sessions, in each session they perform each of the 4 trials previously discussed. During their first visit, they perform the trials without the cooking aid system. The purpose of this first session is to familiarize the participants with the tasks and get a baseline measure of performance. During their second visit, they are provided with a high level explanation of the system and given 5 minutes of "play around" time with the system. This is done with the purpose of resembling the type of interaction someone may have with a newly installed system in their own home. After playing around with the system, participants are asked to explain how they would describe using the system to a friend. This is asked in order to get an initial impression of the participant's mental model of the system. Participants are asked to return for another 4 sessions. The sessions are spaced with approximately 2 days in between each session.

CURRENT RESULTS AND DISCUSSION

There have been a number of observations that have been made so far related to the different use of the cooking aid as a function of age. For example, the use of the system by younger adults has been mostly for verification purposes rather than real-time feedback. Conversely, many of the older adults use the system to provide them with real-time feedback of where they are in the process of completing a specific recipe. It is also worth noting that even the older adults who have relied on the system were not very dependent on it during the condition without distractions. Therefore, the cooking aid, as many automated support systems, appears to be most useful during instances of increased workload.

It has also been observed that older adults re-shape their cooking strategies more often to accommodate to the limitations of the system. For example, some older adults have adopted the strategy of adding about half of one ingredient and moving on to another while the system "catches up." This pattern of results provides some evidence that when useful, older adults are not only willing to use technology and automation but are also willing to change their strategies to ensure that they appropriately rely on it.

Another interesting observation thus far is that older adults are much more hesitant to leave a task unfinished to attend to another task. For example, during interruptions, younger adults will leave the cooking task in the middle of adding an ingredient to attend to the alarm or the stove simulator, while older adults will usually finish adding the ingredient before they attend to the interruption. Overall, participants both younger and older have been able to identify the limitations and capabilities of the system through experience and minimal instruction. The preliminary results of this study suggest that once participants' mental representation of the system is accurate, they will begin to appropriately use the system.

ACKNOWLEDGEMENTS

This research is supported in part by a grant from the National Institutes of Health (National Institute on Aging – Training Grant). It was also supported by the Aware Home Initiative at the Georgia Institute of Technology.

REFERENCES

Cohen, M.S., Parasuraman, R., & Freeman, J.T. (1998). Trust in decision aids: A model and its training implications. *Proceedings of the 1998 Command and Control Research and Technology Symposium.* Washington, DC: CCRP

Dingus, T. A., Mcgehee, D. V., Manakkal, N., Jahns, S. K., Carney, C. & Hankey, J. M. (1997). Human Factors Evaluation of automotive headway maintenance/collision warning devices. *Human Factors*, 39, 216-229.

Lee, J. D., & Moray, N. (1992). Trust, control strategies, and allocation of function in human-machine systems. *Ergonomics*, *35*, 1243-1270.

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors, 46*, 50-80.

McCarley, J. S., Wiegmann, D. A., Wickens, C. D. & Kramer, A. F. (2003). Effects of age on utilization and perceived reliability of an automated decision-making aid for luggage screening. *Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting.* Santa Monica, CA.

Mynatt, E. D., Melenhorst, A. S., Fisk, A. D., & Rogers, W. A. (2004). Aware technologies for aging in place: Understanding user needs and attitudes. *Pervasive Computing*, 3, 36-41

Parasuraman, R. & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, 39, 230-253.

Sanchez, J., Ezer, N., Rogers, W. A. & Fisk A. D. (in press, 2005). Designing collaborative environments for older adults. *11th International Conference on Human-Computer Interaction*, Las Vegas, NV.

Sanchez, J., Fisk, A. D. & Rogers, W. A. (2004). Reliability and age-related effects on trust and reliance of a decision support aid. *Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting* Santa Monica, CA: Human Factors and Ergonomics Society

Wickens, C. D., Lee, J. D., Liu, Y., & Gordon-Becker, S. E. (2004). *An Introduction to Human Factors Engineering*. Upper Saddle River, NJ: Prentice Hall.