A Model of Trust and Reliance of Automation Technology for Older Users

Geoffrey Ho, Liana Maria Kiff, Tom Plocher, Karen Zita Haigh
Honeywell Technology Center
3660 Technology Drive
Minneapolis, MN 55418
geoffrey.ho@honeywell.com

Abstract
To address the growing needs for care for our aging population, both the public and private sectors are turning to advanced technologies to facilitate or automate aspects of caregiving. The user, who is often an older adult, must now appropriately trust and rely upon the technology to perform its task accurately. However, there is reason to believe that older adults may be more susceptible to unreliable technologies. This paper reviews the work on trust and complacency with older adults and examines the cognitive reasons why older adults are more at risk from faulty technologies. Lessons learned from Honeywell’s Independent LifeStyle Assistant™ (I.L.S.A.) are discussed with respect to new high-level requirements for future designs.

1 Introduction
The majority of the western world will experience a doubling of adults over the age of 65, and because older adults require more medical resources than younger individuals, there is a need to accommodate the medical needs of this population. Advanced technologies are promising alternatives to traditional medical care; however, there are drawbacks to technology. Miller et al. [11] describe some of the ethical issues and design goals for designing trustworthy eldercare systems. For instance, the user must feel confident that he or she can learn how to use the device without too much effort. As well, the technology should not result in social isolation, but instead should encourage positive human interactions. Additionally, the user must adequately trust the technology, but also not rely on it to the extent that they become complacent to automation errors. This paper focuses on the specific issues of trust and complacency, which is defined as non-vigilance as result of high trust.

Because of age-related cognitive changes and because seniors tend to be less experienced with computer technology, they may be less able to detect and respond to errors in automated systems. Thus, the errors made by an unreliable device that manages their health may go unnoticed, leading to potentially fatal consequences. This paper reviews the work on trust and complacency in the elderly, provides a new theoretical account of why older adults are more susceptible to automation reliance errors, and discusses how this information can be applied to future designs of automation for elderly caregiving.

2 Previous Work on Trust and Complacency
The research examining trust and complacency of technology amongst older adults is still in its infancy. Only a handful of studies have been published on this issue, likely because technologies supporting senior care are still not widely deployed. However, even among these few studies, several trends have emerged that suggest that the issue is a real concern.

A number of studies now suggest that older individuals trust automated decisions more than younger individuals do. Lee et al. [9] compared older and younger drivers using an Advanced Traveler Information System (ATIS) that presented warning messages to alert drivers of upcoming traffic events. Compliance measures and subjective measures of trust were recorded. They found that older adults were more likely than younger adults to trust the information provided by the ATIS. In a separate study, older adults also trusted an ATIS system that warned of upcoming traffic congestion [2]. In this study, drivers encountered a fork in the road and had to choose the least congested route. The ATIS warned of upcoming congestion at varying degrees of reliability. Again, older individuals trusted the system’s decision more than younger individuals did. Sanchez et al. [15] presented younger and older adults a driving video task in which participants were required to count objects on the roadways while monitoring gauges. A decision support system provided information on the gauges, but at three levels of reliability. Decisions made by the system were always correct (100% condition), correct 80% of the time (80% condition), or only correct 60% of the time (60% condition). They found ratings of trust were more sensitive for older adults between the 80% - 60% conditions.

More recently, Ho [6] presented older and younger individuals with a medication management simulation that operated under different levels of reliability. In the simulation, a computer reminded participants of which medications to take, when to take them, and the dosage. However, when the automated decision failed, it either missed the medication event entirely, or it instructed participants to take an incorrect dosage. He found that older adults not only placed greater trust in the system, they also made more errors in their simulated medication regimen because they relied on the system too much. Specifically, older adults were more likely to miss their medications if the device failed to notify them of a medication event and older adults were also more likely to take the wrong dosage, following the ill advice of the faulty automation. Additionally, although this problem occurred more often in the high reliability condition, it also occurred frequently even when the reliability was poor (67.5% accurate).

Over-reliance on automation for seniors is consistent with findings by Vincenzi and Mouloua [20]. They presented older and younger participants with a modified version of
the Multi-Attribute Battery (MAT) flight simulation task. Because previous work [5, 21] only included one or two subtasks, to increase the demands on workload, the task required participants to perform a tracking task, a system monitoring task, and a fuel management task. The system monitoring task was automated, but participants were still required to ensure that the automation was performing appropriately.

While trust was not examined, they found that under high automation reliability, both age groups were poorer at detecting automation failures. Older adults, likely as a result of the greater workload, were even more complacent than younger adults. That is, they were less likely to notice automation errors and correct for them when they occurred, suggesting that older adults were more reliant on the automation relative to younger users.

These studies highlight a critical issue in technologies designed to be used by older adults. Namely, that older users will tend to trust these devices and, if the systems are unreliable, they may fail to detect and correct for the error. As a result, they will continue to perform an inappropriate action. For instance, if a medication management system incorrectly instructs an older adult to take a medication, the user may be more likely to follow those instructions. Why are older adults more susceptible to automation errors? Several related factors may be involved. Older adults may be less familiar with computer technology and less aware of the potential unreliability. Additionally, cognitive changes related to aging may lead to deficits in interpreting these errors. These cognitive changes may also lead to poorer self-confidence in manual performance and thus the older individual assigns a supervisory role to the computer system, assuming that the system is more knowledgeable [5]. In the next section, the cognitive changes that occur with age will be discussed in relation to trust and reliance on automation.

3 Age-related Cognitive Changes that Affect Automation Reliance

With age, a number of cognitive processes begin to change. These changes occur despite the absence of clinical disease or any environmental mishaps. Most of these changes are minor and many individuals can function normally even in their old-old years. However, age-related cognitive changes are robust effects and may be evident even in the young-old.

The primary cognitive changes include slowing, and deficits in basic attention, memory, and learning, although higher order executive functioning also shows some deficits, including decision-making and reasoning [14]. Several of these processes are necessary when using any form of advanced computer technology. The cognitive change that is probably most obvious to the general public is the difficulty older adults have learning new applications. Several of these processes are also influential in our attitudes towards automation and how we interact with automation. Particularly, attention, working memory, mental workload, decision-making and interpreting stochastic information are discussed as having critical roles in using and relying on automation.

3.1 Attention-Allocation

The failure to attend to the correct information has been implicated as the root cause of complacency [13]. Because automation generally works well, users will learn to trust the automation, and they will direct their attention to sources not under the control of automation. As a result, when an occasional automation error occurs, the user misses the error. While it is well-established that older adults have more difficulties with selective attention [14], basic visual search studies have shown that the attention-allocation strategies used by older and younger users are quite similar [7].

Ho [6] also suggested that poor attention-allocation alone cannot account for complacency effects in older users. In his study, older users had a greater tendency not to monitor and attend to the automation, yet they still made significantly more automation-induced errors. The field study of Honey-well’s Independent LifeStyle Assistant™ I.L.S.A. reported similar monitoring behavior. When first installed, clients monitored I.L.S.A. several times each day. Though there was no requirement to do so, most I.L.S.A. subjects continued to interact with I.L.S.A. daily to check the reports, even after several months of use.

Senders [17] argued that when an operator is monitoring values, he or she develops mental representations of those values. If those values decay from memory, operators will have to resample the data. Those individuals who have poorer working memory will necessarily be required to do more monitoring. Because it is well-documented that older adults show deficits in working memory, it is not surprising that they will also do more monitoring behavior as a result of a greater need to sample. Thus, even though they sample more, older adults are still likely to make automation-induced errors. This finding raises the question, what are other potential causes for automation reliance?

3.2 Working Memory

Deficits in working memory may provide an alternate explanation as to why older adults are more reliant on technology. As mentioned, working memory deficit in older adults is a factor in sampling behavior. Deficits in working memory create two additional problems for older users when faced with automated decisions.

First, without an accurate mental representation of what the appropriate values of the decision should be, older adults are less able to judge whether the automation is making an appropriate decision. For instance, if an individual does not have an accurate mental representation of his or her medication regimen, then they cannot judge the accuracy of directives given to them by automation simply by comparing the contents of the directive to their mental representation.

Second, to determine whether automation is making an appropriate decision, older adults must devote additional cognitive and physical resources to determine whether the system is performing accurately. Furthermore the individual may have an inadequate mental representation of how the automation performs its task, or the information that is being processed, thereby leading to a false assessment of accuracy, either positive or negative. Regardless of the individual’s specific capability to assess the automation, the individual now must create an additional subtask which may involve complex reasoning and problem solving, and results in increased workload.

3.3 Mental Workload

As a result of cognitive changes, older adults generally experience greater workloads for equivalent tasks relative to their
younger counterparts. Although automation is intended to reduce workload, users often do not feel any reduction in workload, likely because the user must still monitor the automation and correct for any errors made by the device. As well, any reduction in workload resulting from automation is negated by the user taking the opportunity to focus on secondary tasks. Thus, when automation errors do occur, it is not surprising that workload exceeds the normal capabilities of an older individual.

Workload does seem to be an important criterion for complacency effects seen in older individuals. For instance, Hardy et al. [5] found no age differences in complacency effects when participants only performed two subtasks in a flight simulation experiment. Similarly, null age complacency effects were also found in a separate single and dual-task study [21]. However, in the two studies where age differences in complacency effects are significant, workload was high in each case.

Recall that Vincenzi and Mouloua [20] deliberately gave participants three subtasks of the MAT in order to create more workload and found, as a result, more complacency for older adults than younger adults. Interestingly though, they did not find any difference in subjective ratings of workload between younger and older participants. In contrast, Ho [6] found that older adults were more reliant on automation and found that older users were subjectively rating workload higher than younger adults. Thus, based on these studies, it appears that high workload is a necessary condition for age differences in automation reliance.

It is important to note that automation may also be guilty of decreasing an elderly person’s mental acuity by inappropriately reducing workload. The I.L.S.A. medication compliance system used a strategy that “rewarded” test subjects for not relying on the automation by supressing the automated telephone reminder when it was not required. [4] The I.L.S.A. medication compliance system reminded test subjects to take their medication only after it was clear they had forgotten by employing a configurable window around each dose event. I.L.S.A. subjects were monitored for adherence behavior for two weeks after the automation interventions had been removed. Though the sample size is too small to be statistically significant, the behavior trends showed a continuation of self-reliance for the test subjects who had been responsive to the described “reward” system.

3.4 Decision-making

Decision-making also changes with age. While some of these age differences may reflect increased cautiousness and wisdom with age, decisions are affected by other cognitive factors such as workload. Particularly under time pressure, people tend to use more heuristic processing. Heuristics are simple cognitive solutions that work well but may lead to inappropriate actions if more complex reasoning is required. Automation can be used like a heuristic because following automated directives provides an easy solution without cognitive overhead [19].

In conditions of high workload, users of automation may not consider the inaccuracy of automated information because it adds even more workload. Instead, a heuristic approach is used; simply follow the automated advice because it generally works well. This bias to use automation despite its occasional unreliability may or may not be related to trust and complacency, but may simply be the best alternative to deal with a high workload situation. This behavior reflects how users will often select the most “satisficing” decision, not the most accurate [18].

Studies in decision-making in older adults indicate that they take advantage of heuristic processing more often than younger adults [1]. Chen and Sun [1] examined financial decision-making by having younger and older participants conduct a simulated yard sale. Subjects had to agree or disagree to offers for merchandise and their goal was to make as much money as possible. Strategies for selling were either a single-deal strategy whereby the first offer was rejected, or a multiple-deal strategy whereby the first offer was accepted. The multiple-deal strategy was considered more complex because it required one to compare the offer with previous offers in working memory to obtain the largest profit. The results indicated that older users more often used the single-deal strategy and the authors suggested that older people adopt a less cognitively demanding decision process.

3.5 Interpreting Stochastic Information

Older adults also have difficulty interpreting the frequency of errors from machines. In a series of studies, Sanford and Maule [10; 16] had older and younger participants perform tasks using devices that varied in their reliability, which illustrated that older users did not change their monitoring strategies to reflect the differences in reliability. In one study [10], they had participants perform a fault detection task on three machines, each with a different level of reliability. Sixty percent of the faults occurred on the first machine, 30% occurred on the second machine, and 10% on the third machine. A meter indicated when a fault was occurring, but the participants had to determine which of the three machines were faulty. According to the authors, the best strategy would be to monitor the least reliable machine first. They found that with experience, younger adults began to monitor the least reliable machine more frequently, but older adults did not adopt this strategy, suggesting that older adults are less sensitive to machine behavior, resulting in less efficient strategies for monitoring. The authors argued that older individuals may be less apt at interpreting stochastic information of the device and as a result, they perform fewer adjustments to their strategies to accommodate for the errors. This finding suggests that older users will be less apt to notice or adjust their usage of automation because they fail to recognize the frequency of faults. If a machine were mostly accurate, older users would adopt it.

4 A Holistic View of Automation Reliance

No unique cognitive process can be identified as a root cause of automation reliance, but rather all of these cognitive processes are dependent. Figure 1, taken from Ho [6], depicts a user performing a dual-task, one with the help of automation. It illustrates how these cognitive dependencies interact and how the combination of these factors influences automation reliance. Several important points of this figure should be highlighted.

First, note that this process is cyclical. Trust is only one, albeit critical factor, that affects automation usage and there are many factors that can influence trust itself. For example, trust can influence how one allocates attention. Memory and attention can affect workload, which in turn affects decisions to use automation. Additionally, the ability to interpret the frequency of errors will affect trust and bias decisions.
Second, note that when the user decides to verify automation accuracy, a tertiary task is created. It does not assume that monitoring automation is a simple visual task. Instead, it may require complex reasoning and problem solving to determine the veracity of automated decisions. Because this tertiary task is added, efficiency decreases to a point that is beyond not using the automation at all.

Third, while not explicitly shown, all of the cognitive abilities mentioned in the previous section act to influence this entire process. Trust, along with several other factors affects the attention allocation strategy and decision bias. Attention strategies and working memory will affect the acquisition of data. Working memory and mental workload will affect the decision to verify automation accuracy. Additionally, a high workload generally leads people to follow the top route, the bias to use automation. Finally, the outcome of the decision will influence future decisions, however in the case of older adults, the ability to use this information in a stochastic fashion may be less efficient.

This process view of using automation provides a holistic understanding of why older adults may use automation differently. The dependencies of human cognition make it difficult to pinpoint any one reason why older adults may use automation differently. However, by identifying the various causes and their relationships, we can begin to better understand the constraints of designing systems that automate caregiving tasks so as to provide sufficient support while encouraging appropriate levels of trust and minimizing over-reliance.

5 Lessons Learned from I.L.S.A. and Future Research

Honeywell’s Independent LifeStyle Assistant™ I.L.S.A. is an autonomous, context-aware monitoring and caregiving smart-home system intended to provide support to older adult clients and their caregivers. Critical needs identified included medication management, toileting, mobility, safety, and caregiver burnout. For a more detailed description of I.L.S.A., please refer to Haigh et al. [3; 4]. From March through July of 2003, a field study of I.L.S.A. was deployed in eleven homes of older individuals. A number of observations from this field study illustrate trust and reliance issues.

5.1 Perceived Reliability

From a system perspective, reliability is at the heart of trust and reliance. After all, a system that unreliable will not be trusted, relied upon or even used. Findings from the I.L.S.A. field study were illuminating as to the complexity of accurately performing monitoring and caregiving using technology. For instance, medication reminders were perceived to be inaccurate, even when the system was behaving as designed. One user claimed that I.L.S.A. reminded her to take a medication when she had decided to take the medication half an hour earlier, one minute outside the “grace period” around her dose time. Another user also commented on I.L.S.A. not recognizing a change in her usage pattern, a feature developed in the lab but not provided in the elders’
homes. This perceived unreliability of the system resulted in at least one of the clients reporting that she didn’t trust I.L.S.A.

Unreliability, real or perceived, doesn’t simply lead to distrust, but adds demands to working memory and as a result leads to increased workload. The negative effect associated with the system adds to feelings distrust of the system. Some will reject the automation and the additional workload after the first or second failure. Unfortunately, improving reliability can be expensive and even then, 100% accuracy is unlikely. For example, in our occupancy data, single sensors were inaccurate at determining occupancy; while multiple sensors were much more accurate, they did not achieve 100%. Thus, the costs of gathering redundant evidence must be weighed carefully.

These results raise the question: “to what extent should automation for the elderly be accurate?” The fear that most human factors researchers have is that a system that approaches 100% accuracy will lead to complacent behavior. Yet, dealing with rare instances of complacency may be the lesser of two evils when the user is an older adult. Ho [6] showed that when under high workload, older adults tend to over-trust and are subject to complacency-like errors even when system reliability is relatively low (65 - 90%). Thus, extending reliability as close to 100% accuracy may be optimal for older users. More research is needed and studies investigating countermeasures of complacency such as adaptive automation has yet to be attempted using older participants. This avenue of research may be valuable in determining optimal reliability of future systems.

Finally, achieving perceived reliability, or gaining trust, is a matter of setting appropriate expectations. Given a particular set of sensor evidence, we may be able to determine with 100% accuracy that no one is moving in the home. However, we may be only 50% confident that the lack of motion is due to the absence of an occupant, and still less confident that the lack of motion constitutes an emergency. Clearly, setting the right expectations for automation is equally important to the elderly user and their caregivers. In the medication example noted earlier, for example, a communication method that allowed the user to accurately perceive the events as they occurred may have assured the elderly client that the system was performing as designed.

5.2 Trust and User Interfaces
The I.L.S.A. field studies also produced a great deal of knowledge regarding methods of allowing the elderly to interact with automation. I.L.S.A. had two modes of communication with elderly clients; Honeywell’s WebPAD™, a touch-screen interaction device, and the telephone. Both interaction modes concentrated on one way communication from I.L.S.A. to the client. This behavior was favored to allow for entirely passive or non-interactive clients.

Older clients did not like the recorded female voice that represented I.L.S.A. in the telephone reminders. Many of the older participants in the I.L.S.A. field studies commented that a friendlier, more polite version of I.L.S.A. was desired. For instance, when one user was asked what she would say if she were I.L.S.A., her response was “A cheerful good morning.” Evidently, I.L.S.A.’s cheerful, but short, “Hello!” was not enough to convey a sunny disposition. The text of I.L.S.A.’s interface was certainly polite, but the format and imposition of a recorded interaction requires extra care to overcome the negative aspects of being addressed by a machine. It may be due to the inability to interact effectively with the telephone system that users described it as “computer generated.”

Unfortunately, technologies driving interactive speech interfaces are generally not sophisticated enough to handle complex interactions with older adults. Koester [8] and Haigh et al. [4] provide additional evidence that speech interaction is unlikely to be effective in this domain. Among other issues, they point out that there is a significant “cognitive cost” to use speech recognition systems; the elders most suited to use an I.L.S.A.-like system are unlikely to be capable of meeting this challenge. Furthermore, many elderly people experience significant change in their speech patterns and cognitive capacity throughout the day, delivering special challenges to both user and system.

An interactive speech interface is appealing because it provides a natural avenue for affective computing. Personification of interfaces may elicit greater feelings of trust [1]. Affect is becoming increasingly acknowledged as a major influence of trust. For instance, Parasuraman and Miller [12] found that “polite” voice interfaces were rated to be more accurate than “impolite” voices even though the opposite was in fact true. The effects of affects, personified interfaces may have even stronger influences on older subjects. If the elderly are already prone to trust automation and are particularly sensitive to emotional responses, then they may be more influenced by affective interfaces. Though endowed with a female name through the I.L.S.A. acronym, the I.L.S.A. interface did not suggest personification in any respect other than the female vocalizations used to deliver medication reminders. However, many of the subjects consistently referred to I.L.S.A. as “she” throughout the study.

Though special attention was paid to keep interactions simple, keep workload low, and support memory rather than tax it, clients still had trouble with the telephone-based delivery of medication reminders. I.L.S.A.’s reminder system asked the clients to confirm receipt and comprehension of the medication reminder by pressing either “1” or “2.” Given the prevalence of telephones with the keypad integrated in the receiver/handset, clients had to shift context, look at the keypad, remember what they were asked to do, and then return to listening. Many reported annoyance and some unease with this task as it taxed working memory [4].

Though the telephone interface presented challenges with the simplest task, test subjects were generally confident with the touch-screen device and desired a more interactive interface [4]. Event-based systems such as graphical user interfaces reduce workload and compensate for poor working memory by allowing users to procede at their own pace, without feeling pressured to perform.

Finally, a further example from I.L.S.A. illustrates how users can interact with a system to create entirely new reliability issues. I.L.S.A.’s medication adherence system encouraged clients to remember medication events on their own and avoid receiving the telephone reminder, which they disliked intensely. Several clients learned that they could fool the system by accessing their medication container on time, while choosing not to take the medication for a variety of legitimate reasons. However, one of these same clients also reported that she used I.L.S.A.’s adherence report to remember if she’d taken her pill already. Clearly, it would be easy for her to trick herself as well as the system, and miss a pill she had only “pretended” to take earlier. Thus, it is important to consider all potential uses and mis-uses of your
system when assessing it for trust and reliability.

6 Conclusion

Trust and reliance on automation will be a particularly important issue if technology is to serve as a caregiving tool for the elderly. The data from Honeywell’s I.L.S.A. project clearly demonstrated the distrust experienced by our participants as a result of real and perceived unreliability. While high unreliability is obviously unacceptable, as technology for caregiving advances, issues of trust and over-reliance related to near-perfect reliability also need to be addressed.

A number of cognitive processes influence trust and reliance on automation and these processes provide designers with high level requirements and constraints for future designs.

- Compensate for limited working memory of older adults
- Limit the number of reminders and other interruptions that add to workload and frustration
- Design feedback to enable accurate interpretation of automation outcomes
- Aid the user to set reasonable expectations of automation reliability

This paper offers a holistic explanation of why older adults may be more susceptible to automation errors. While this model of automation reliance is based on decades of research on older adult cognition, only a few studies have investigated automation reliance using senior participants; even fewer have conducted long-term field tests, which can provide more accurate assessment of complacency. Further research is required to improve the design of automated caregiving systems and their interactions with all users.

References


