THEORETICAL ISSUES IN ERGONOMICS SCIENCE, 2016
http://dx.doi.org/10.1080/1463922X.2016.1175523

Taylor & Francis
Taylor & Francis Group

# Does the domain of technology impact user trust? Investigating trust in automation across different consumer-oriented domains in young adults, military, and older adults

Richard Pak [a], Ericka Rovira[b], Anne Collins McLaughlin[c] and Natalee Baldwin[a]

[a]Department of Psychology, Clemson University, Clemson, SC, USA; [b]Department of Behavioral Sciences & Leadership, US Military Academy, West Point, NY, USA; [c]Department of Psychology, North Carolina State University, Raleigh, NC, USA

**ABSTRACT**
Trust has been shown to be a determinant of automation usage and reliance. Thus, understanding the factors that affect trust in automation has been a focus of much research. Despite the increased appearance of automation in consumer-oriented domains, the majority of research examining human-automation trust has occurred in highly specialised domains (e.g. flight management, military) and with specific user groups. We investigated trust in technology across three different groups (young adults, military, and older adults), four domains (consumer electronics, banking, transportation, and health), two stages of automation (information and decision automation), and two levels of automation reliability (low and high). Our findings suggest that trust varies on an interaction of domain of technology, reliability, stage, and user group.

**ARTICLE HISTORY**
Received 5 December 2015
Accepted 3 April 2016

**KEYWORDS**
Trust in automation; domain; aging; complacency; military; technology; stages of automation; levels of automation



## Relevance to human factors/ergonomics theory

The results of this study demonstrate the large effects associated with different domains and groups in trust in automation. The findings highlight the importance of representative design in human factors research and the potential limits in translating previous results to consumer-oriented technology (Fisk and Kirlik 1996; Czaja and Sharit 2003).

## Introduction

Historically, technological automation was relegated to specific use cases with highly trained users (e.g. complex flight automation, process-control). However, users of all types are being exposed to more automation in many consumer-oriented contexts. For example, digital cameras have automatic scene detection, banks allow automated bill payments, and cars assist with parallel parking. The extent to which users trust automation is a key determinant of use and reliance (Lee and Moray 1992, 1994; Muir and Moray 1996). When

**CONTACT** Richard Pak richpak@clemson.edu

users' confidence in carrying out a task exceeds their trust in the automation's ability to do so, they will often forgo the use of automation (Lee and Moray 1994). Thus, understanding the system-related factors as well as user characteristics that affect trust in automation has been a focus of research (for a review, see Hoff and Bashir 2015).

The majority of research examining trust in automation has typically used highly specialised domains such as supervisory or process control tasks and with highly trained populations (Hoff and Bashir 2015). An example is Lee and Moray's (1992) seminal study examining trust in automation. The study was vital in understanding the factors that affected changes in trust in automation; however, the idiosyncratic nature of a process-control task might not generalise to more consumer-oriented automation (as Petrinovich 1989 suggested). For example, consumer-oriented automation tends not to require continuous and intense monitoring, or have severe consequences (job loss or injury).

The possibility that automation may not be trusted equivalently by users depending on the context of use was first noted by Inagaki (2006) who argued that applications of automation within transportation systems might need to be tailored to the specific mode of transport. His rationale was that each mode of transportation came with differences in operator (different levels of training required for drivers versus pilots), and time criticality or severity (with greater time criticality for ground transportation compared to aviation). These differences, he argued, should alter the preferred type and level of automation and the operator's trust in the system. He concluded it was necessary to carry out research for each transportation mode (driving, aviation). To be sure, there are examples of studies that have either sampled more consumer-oriented context (e.g. Ho, Wheatley, and Scialfa 2005; Sauer and Ruttinger 2007) or have used participant populations that matched the intended population (e.g. Calhoun, Draper, Ruff 2009; Mosier and Fischer 2012) but these studies have not examined possible differences in trust due to the domain within the same study.

In addition to the evidence presented above, there are several models or frameworks of trust in the automation literature that led us to expect that domain might moderate the level of trust placed in automation (e.g. Hoff and Bashir 2015; Ho, Kiff, Plocher, and Haigh 2005; Lee and See 2004; Dzindolet et al. 2001; Parasuraman and Riley 1997). Hoff and Bashir's recent framework is most detailed in elaborating on the possible role of domain on trust. They identified two broad factors that affect trust in automation: external factors (those related to the task or the automated system such as context or task type) expected to affect situational trust and internal factors (those related to the participant or group such as age or culture) expected to affect dispositional trust. However, these models only tacitly suggest the role of domain of automation on trust, often referring to 'context' which we interpret to mean within a single domain (e.g. variations of the system or task context within the domain of transportation; Inagaki 2006). Thus, the nature of the effect of domain, which can vary independent of context, and how it interacts with other external or internal factors, is an open empirical question. A brief review of relevant internal and external trust factors that may be relevant to domain follows.

## *External factors that influence situational trust in automation*

### *Automation domain*

Hoff and Bashir (2015) found that nearly 80% of trust in automation studies (the top four categories of Table 1) was in domains that could be considered military/security/

**Table 1.** Examples of automated systems (from Hoff and Bashir 2015).

| Type of system | Number of studies ($N = 127$) | Percentage of total |
|---|---|---|
| **Combat identification aid** | **31** | **24.4** |
| **General decision aid** | **25** | **19.7** |
| **Fault management/task monitoring aid** | **24** | **18.9** |
| **Automated weapons detector (luggage)** | **11** | **8.7** |
| Target identification aid (noncombat) | 9 | 7.1 |
| Collision warning system | 9 | 7.1 |
| Route-planning system | 7 | 5.5 |
| Other | 11 | 8.7 |



industrial. The overrepresentation of mission-critical domains in automation research perhaps reflects actual and predominant uses of early automation in military/security, aviation, and industrial settings (e.g. Fitts 1951; Warren 1956). These domains involve automation that alleviated high vigilance requirements to maintain operations. Hoff and Bashir (2015) concluded that there was a paucity of research utilising consumer-oriented technology. Due to the overrepresentation of high-criticality industrial or military domains in human automation interaction literature, it is unknown the extent to which research in one specific domain, typified by high consequence and severity, can be applied to the design of more consumer-oriented systems (generalisability) or integrated with other results to evaluate theories of behaviour (Rogers, Pak, and Fisk 2007).

### *Stage of automation*

In addition to classifying the domains of automation that have been studied, another, more psychologically relevant way of classifying the automation literature is by describing what is being automated and by how much. This characteristic is termed *degree of automation* (DOA; Wickens et al. 2010; Parasuraman, Sheridan, and Wickens 2000) with higher DOA supporting *higher levels* (Sheridan and Verplank 1978) and *later/higher stages* of human information processing (e.g. decision-making rather than perception). Sheridan and Verplank (1978) described levels of automation that ranged from fully manual to fully autonomous. Parasuraman, Sheridan, and Wickens described four *stages* of automation, in order of increasing support, but the boundary between information automation (stages 1−2) and decision automation (stages 3 and 4) represented a critical distinction as the detrimental effects of unreliable automation worsened above the boundary (Onnasch et al. 2013). In information automation, the operator is kept 'in the loop' as they are responsible for the decision-making aspects of the task and thus are expected to have better situation awareness than with decision automation (Endsley and Kiris 1995). For these reasons, higher stages of automation tend to be associated with increased performance decrements because operators are caught over trusting the automation without sufficient time and information to recover (Crocoll and Coury 1990; Sarter and Schroeder 2001; Rovira, McGarry, and Parasuraman 2007).

In their review of the literature, Hoff and Bashir (2015) found that about 75% of the studies in their pool examined the effect of relatively higher stages of automation (decision selection and action implementation) on trust (Table 2). The authors surmised that a likely reason for this trend to study higher stages of automation versus lower was that most decision selection automation, by virtue of its complexity and opaqueness, might require more trust from users. Another reason may be that it was a natural consequence

**Table 2.** Categories of automation used in eligible studies (from Hoff and Bashir 2015; reordered by increasing stage).

| Category of automation (automation stage) | Number of studies ($N = 127$) | Percentage of total |
|---|---|---|
| 1. Information acquisition | 2 | 1.6 |
| 2. Information analysis | 25 | 19.7 |
| **3.** Decision selection | **95** | **74.8** |
| 4. Action implementation | 5 | 3.9 |

of the overrepresentation of mission-critical domains where automation that saves time in decision-making was valued.

## *Internal factors that influence dispositional trust in automation*

### *Culture*

In addition to factors related to the system, characteristics of the user are also expected to influence trust in automation. Hoff and Bashir (2015) identified several internal factors that influenced dispositional trust in automation such as culture, age, gender, and personality traits. Of the listed factors, Hoff and Bashir concluded that research specifically on the influence of age and culture on trust in automation has been underrepresented. Although Hoff and Bashir's conception of culture was in the international sense, effects of organisational culture are also underrepresented in the research and warrant further examination.

One highly distinct organisational culture is the military. Culture can be defined as the learned and shared meanings, ideas, and symbols that distinguish a group or category of people (Hofstede 1991; Soeters, Winslow, and Weibull 2006). Military culture is characterised by a hierarchy with a strong social order that follows and imposes an established set of rules and regulations, a high level of discipline, and compliance of rules and acceptance of authority (Soeters, Winslow, and Weibull 2006; Hollands and Neyedli 2011).

Kennedy, Sibley, and Coyne (2015) proposed that military training and indoctrination might affect expectations of and behaviours with automation such that soldiers are likely to treat automation as a teammate. The mentality of automation as teammate may lead the soldier to set initially unreasonably high expectations of the automation's capability and its awareness of the situation. However, studies (e.g. Wang, Jamieson, and Hollands 2009; Dzindolet et al. 2001) which were meant to generalise to a military population often used college studies for convenience which may have eliminated understanding of the role of organisational culture on automation trust and reliance. No current studies have empirically demonstrated that organisational culture, such as military culture, has an effect on trust in automation.

### *Age*

Another understudied factor affecting dispositional trust in the automation literature is age. Hoff and Bashir (2015) identified 8 of 127 studies examining the effects of aging on trust in automation. One of the key findings seemed to be that older adults tended to comply with and rely on automation more than younger adults (McBride, Rogers, and Fisk 2011); that is, they are more complacent (over trusting) with automated systems than other age groups (Ho, Wheatley, and Scialfa 2005; Ho, Kiff, Plocher, and Haigh 2005). A major hypothesis to explain older adults' tendency for over-reliance and complacency is that age-related cognitive changes in working memory make it more difficult to detect automation faults (Ho, Wheatley, and Scialfa 2005).

### *Research question and hypotheses*

The purpose of the current study was to fill in the following gaps in the literature on trust perceptions of automation by investigating:

(1) The extent to which domain (with an emphasis on more commonly encountered, consumer-oriented domains) differentially affects trust in automation.
(2) The extent to which automation stage and reliability moderates trust within consumer-oriented domains.
(3) How organisational culture (specifically military) and aging affects trust in consumer-oriented technology of varying domains and stages as compared to civilian college students.



Understanding general principles of behaviour (i.e. factors that affect trust in automation) requires researchers to sample a wide range of user characteristics and representative (consumer-oriented) situations (representative design; Hammond 1998). It is assumed that the current body of trust in automation literature, carried out in one context with a specific set of users, will be generalisable to other contexts and users as long as the conceptual aspects (e.g. reliability, stage) of automation are identical. To evaluate this implied generalisability, we decoupled automation domain from characteristics inherent to the automation (e.g. reliability, stage) and measured trust in three different groups of users that systematically varied in organisational culture and age group. Knowing the limits of generalisability is useful for theory advancement but also achieves the practical goal of accurate prediction of trust in automation with consumer-oriented technology or with lay users.

This study was partially exploratory to examine the moderating roles of domain and group on trust. Thus, although we had specific hypotheses guided by existing literature, we expected these 'main effects' to be moderated by domain and group. We hypothesised that, consistent with the literature, (1) across all participant groups and domains of automation, participants would have higher trust in more reliable automation (Parasuraman, Molloy, and Singh 1993; Dzindolet et al. 2003; Lee and Moray 1992). (2) We hypothesised that older adults would have higher trust ratings than younger adults across all domains (Ho, Wheatley, and Scialfa 2005; Pak et al. 2012; Pak, McLaughlin, and Bass 2014). (3) We hypothesised trust differences between civilian and military such that military participants would have higher trust judgments overall and their trust would be more resilient to automation failure (i.e. less trust decrement due to reliability). This would be due to the intensive training and general increased discipline and adherence to rules and protocols which may require the use of automation (Kennedy, Sibley, and Coyne 2015).

Regarding the effects of domain on trust, (4) we expected that more familiar forms of automation would be more trusted than less familiar domains for younger adults (Muir and Moray 1996) and older adults (Melenhorst, Rogers, and Bouwhuis 2006). Of the four domains used in this study, the general order of decreasing familiarity based on surveys of technology usage (Olson et al. 2011) might be consumer electronics,

**Table 3.** Participant demographics.

| | Cadets ($N = 53$) | | Students ($N = 59$) | | Older adults ($N = 36$) | |
|---|---|---|---|---|---|---|
| | *M* | SD | *M* | SD | *M* | SD |
| Age | 19.5 | 1.4 | 18.8 | 1.0 | 70.5 | 4.0 |
| CPRS* | 51.6 | 5.2 | 51.5 | 3.6 | 50.6 | 5.3 |

*Scores could range from 16 indicating low complacency potential to 80 indicating high complacency potential (Singh, Molloy, and Parasuraman 1993). There was no significant difference in CPRS between the groups ($p = 0.55$).

banking, health, and transportation. Here, we also expected an interaction between domain and age since technological familiarity is related to age (Olson et al. 2011) with more familiar technologies for each age group leading to higher trust. Finally, consistent with previous literature we expected to find that (5) decision automation would be more trusted than information automation (Rovira, McGarry, and Parasuraman 2007).

## Method

### *Participants*

An a priori power analysis using Faul and Erdfelder's (1992) Gpower showed that a minimum of 126 participants were required for enough power to detect an effect size of 0.2 (a power level of .8 and alpha at 0.05). A total of 146 participants (72 female) were surveyed. Table 3 illustrates the number of US Military Academy cadets, civilian undergraduate students (referred to as students), and civilian older adults (referred to as older adults). Cadets and students received extra course credit, while older adults received $10 for their participation. The older adults were community-dwelling and independent-living (i.e. did not reside in a care facility). This study was approved by the US Military Academy and Clemson University's IRB.

### *Materials*

The study was administered remotely via the Qualtrics web-based survey platform. In addition to the factorial survey, participants also completed the Complacency Potential Rating Scale (CPRS; Singh, Molloy, and Parasuraman 1993), a 16-item scale designed to measure complacency towards common types of automation (e.g. automated teller machines). In the CPRS, participants responded to the extent they agreed with statements about automation on a Likert scale of 1−5. The CPRS score was a sum of these responses and ranged from 16 (low complacency potential) to 80 (high complacency potential).

#### *Automation scenarios*

Factorial surveys gather subjective assessments by using scenarios; they are particularly useful when assessing how experimental manipulations affect subjective perceptions, such as trust (Rossi and Anderson 1982). Additionally, this approach has also been used in prior automation research (e.g., Endsley and Kiris 1995; Mosier and Fischer 2012; Pak, McLaughlin, and Bass 2014). The four domains examined in this study represented a wide range of consumer-oriented technology that would be easily relatable to younger and older adults (in contrast to industrial, security, or military automation). The domains selected also allowed for realistic variability of automation reliability and stage of

**Table 4.** Younger and older adults' familiarity with technology by domain.

| | Younger adults | | Older adults | |
|---|---|---|---|---|
| Domain | *M* | SD | *M* | SD |
| Consumer | 1.26 | 0.47 | 1.08 | 0.60 |
| Banking | 0.89 | 0.65 | 0.73 | 0.70 |
| Transportation | 0.61 | 0.58 | 0.52 | 0.60 |
| Health | 0.35 | 0.45 | 0.47 | 0.56 |

Note: Values are means of familiarity with technologies within each category (0 indicating unfamiliarity to 2 indicating high familiarity) as reported by Olson et al. (2011).
Source: Adapted from Olson et al. 2011.

automation. Because participants did not interact with actual automation but instead were asked to think about the specific scenarios, it was important that the examples of automation were likely to be encountered by the average person so they could form opinions. Table 4 illustrates the relative familiarity of the four categories of technologies among younger and older adults.

Each scenario included three factors: automation domain (consumer electronics, banking, transportation, and health), reliability of the automation (low, high), and automation stage (information, decision). The factorial combinations of the three manipulated factors resulted in 16 unique scenarios. A sample scenario representing *high reliability*, *decision* automation in the *health* domain is presented below:

> Jean was rushed to the hospital and found out she needed a **pacemaker**. Her doctor told her the pacemaker is designed to **automatically adjust to her activity level**. For example, if she is reading a book **it will produce a lower heart rate and when she is exercising it will produce a higher heart rate**. **Six months after the surgery Jean has now resumed her normal activities without any issues.**

Key sections that identify this scenario as health, decision automation, and reliable are bolded for illustrative purposes. It is in the health domain because of the use of a pacemaker. It is an example of decision automation because it is carrying out an action (changing heart rate; action implementation) without the user's intervention and it is of high reliability because the automation performed successfully. An example of *low reliability*, *information* automation in the *consumer electronics* domain is presented below:

> Jack owns a local ranch and prides himself on how well he takes care of his horses. Jack keeps a **weather radio** in his house to alert him of storms so he can get his horses into the barn before the storm arrives. **The radio is programmed to sound an alarm and give storm details for bad weather within 15 miles**. Jack's weather radio recently **alerted him of a storm over 500 miles away**. Because Jack can no longer anticipate severe weather, he cannot move his horses into the barn in advance.

The weather radio in this example was in the consumer electronics domain because it is intended for everyday use; information automation because it serves a simple alerting function (information acquisition) and not carrying out an action. It is low reliability because it is alerting to a distant storm, which is useless to the user.

Flesch–Kincaid readability statistics (Kincaid, et al. 1975) showed that the mean grade level for the scenarios was 8.6. All scenarios were pilot tested to ensure that each manipulation (domain, reliability, and automation stage) was noticeable. In the pilot test, undergraduate participants read each scenario, rate the reliability of the automation presented, name the domain, and judge whether it seemed to be a lower or higher form of

**Table 5.** Factors manipulated and examples of vignettes.

| Consumer electronics | | Banking | | Transportation | | Health | |
|---|---|---|---|---|---|---|---|
| Information automation | Decision automation | Information automation | Decision automation | Information automation | Decision automation | Information automation | Decision automation |
| *Lower reliability* | | | | | | | |
| Word processing spell check | Television digital video recorder | Online banking checking account balance | Online banking automatic bill pay | In-vehicle blind spot alert | Collision avoidance automated braking | Smartphone enabled blood pressure cuff | Pacemaker |
| *Higher reliability* | | | | | | | |
| Weather radio | Word processing auto-correct | Automated teller machine | Online banking automatic transfers to savings | Speed limit alerting | Personal GPS navigation | Blood glucose meter | Insulin pump |



automation. Pilot participants were accurately able to state the domain and relative reliability (low or high) and relative stage of automation (lower stage information or higher stage decision) as intended. Table 5 illustrates the combinations of factors and the example technology used in the scenarios. While the severity or cost of error (Ezer, Fisk, and Rogers 2008) of the automation failure was not explicitly manipulated in the scenarios, it was balanced by having equal numbers of relatively low and high severity domains. Transportation and health were considered higher severity as automation failure could result in injury or death while banking and consumer electronics were considered lower severity as a failure was unlikely to result in injury.

Below each scenario was a series of questions. First, participants were asked, 'what technology is discussed in this story?' This question was used as a manipulation check and all participants were able to articulate the technology. The second question, modelled on Lee and Moray's (1994) trust questionnaire, was, '[character name] should trust the technology portrayed in this story' followed by a 7-point Likert scale ranging from strongly disagree to strongly agree.

## Design and procedure

The study was a 3 (group: students, cadets, older adults) × 4 (domain: consumer electronics, banking, transportation, health) × 2 (stage of automation: information, decision) × 2 (reliability: low, high) within-subjects design with each participant exposed to every domain, stage of automation, and reliability. For each scenario, participants reported their trust in the automation on a Likert scale (ranging from 1 to 7).

Participants were sent a link to complete the survey. They were instructed to complete the survey in one sitting and to avoid taking breaks (web log time entries confirmed this). The 16 scenarios were then presented in a random order for each participant one-at-a-time. Participants completed the experiment remotely taking as much time as they needed. After judging all 16 scenarios, participants completed CPRS.

## Results and discussion

The trust measurements were subjected to a 3 (participant group: students, cadets, older adults) × 4 (domain: consumer electronics, banking, transportation, health) × 2 (stage of

**Table 6.** ANOVA table.

| Sources | SS | df | MS | *F* | *P* | $\eta_p^2$ |
|---|---|---|---|---|---|---|
| | | | Between subjects | | | |
| User group (G) | 63.37 | 2 | 31.69 | 5.19 | 0.01 | 0.07 |
| *S* within-group error | 872.33 | 143 | 6.10 | | | |
| | | | Within subjects | | | |
| **Automation reliability (AR)** | 242.31 | 1 | 242.31 | 62.42 | 0.00 | **0.30**** |
| **Automation stage (AS)** | 218.26 | 1 | 218.26 | 101.79 | 0.00 | **0.42**** |
| **Domain (D)** | 688.48 | 1.91* | 360.30 | 83.24 | 0.00 | **0.36**** |
| **D × G** | 377.29 | 6 | 62.88 | 22.81 | 0.00 | **0.24**** |
| **AR × G** | 737.40 | 2 | 368.70 | 94.98 | 0.00 | **0.57**** |
| **AS × G** | 863.22 | 2 | 431.61 | 201.28 | 0.00 | **0.74**** |
| D × AR | 65.40 | 3 | 21.80 | 17.78 | 0.00 | 0.11 |
| D × AS | 31.25 | 2.64* | 11.86 | 7.12 | 0.00 | 0.05 |
| AR × AS | 17.30 | 1 | 17.30 | 13.09 | 0.00 | 0.08 |
| D × AR × G | 20.51 | 6 | 3.42 | 2.79 | 0.01 | 0.04 |
| **D × AS × G** | **104.55** | **6** | **17.43** | **11.91** | **0.00** | **0.14**** |
| AR × AS × G | 9.77 | 2 | 4.88 | 3.70 | 0.03 | 0.05 |
| D × AR × AS | 16.47 | 3 | 5.49 | 4.44 | 0.00 | 0.03 |
| D × AR × AS × G | 32.35 | 6 | 5.39 | 4.36 | 0.00 | 0.06 |
| D × AR within-group error | 525.95 | 429 | 1.23 | | | |
| D × AS within-group error | 627.78 | 429 | 1.46 | | | |
| AR x AS within-group error | 188.99 | 143 | 17.3 | | | |
| D × AR × AS within-group error | 530.69 | 429 | 1.24 | | | |

Note: *S* = subjects.
*Mauchly's test indicated a violation of sphericity so Greenhouse-Geisser corrected values are reported.
**an effect size (variance accounted for) greater than 0.14 and $p < 0.001$.

automation: lower, higher) × 2 (reliability: lower, higher) repeated measures ANOVA. The results of the ANOVA are shown in Table 6. A commonly encountered problem with the use of ANOVA in studies with a large number of participants is a high number of significant *p*-values. *P*-values do not indicate the size of the effect so for parsimony, our analysis strategy was twofold. We first chose to focus our follow-up analyses and discussion on our a priori hypotheses. Next, we explored the significant two- and three-way interactions that were *practically* significant; that is, those significant results ($p < .001$) that also had effect size (partial eta squared) greater than 0.14. Partial eta squared ($\eta_p^2$) indicates the unique variance accounted for in trust by the independent variable after controlling for influence of other variables. Published guidelines indicate that a threshold of $R^2 = 0.14$, or 14% unique variance accounted for, is considered a *large* effect size (Cohen 1988; Richardson 2011). As a point of comparison in the human factors literature, Hancock, et al. (2011) conducted a meta-analysis of published studies examining trust with robots (a form of automation) and found a *medium* effect size of $r = .26$ ($R^2 = 0.06$), which corresponded to 6% variance in trust due to global factors such as reliability and appearance.

## Main effects

Using this effect size criteria, there were significant main effects of automation reliability, group, domain, and automation stage on trust perceptions (Table 6). Supporting hypothesis 1, trust was greater when automation was more reliable ($M = 5.1$, $SD = 1.8$) than less reliable ($M = 4.3$, SD $= 1.9$). Supporting hypothesis 2, older adults' trust ($M = 5.0$, SD $=$ 1.8) was significantly higher than students ($M = 4.6$, SD $= 1.9$). Contrary to our third hypothesis, cadets and students did not significantly differ on trust. Finally, information

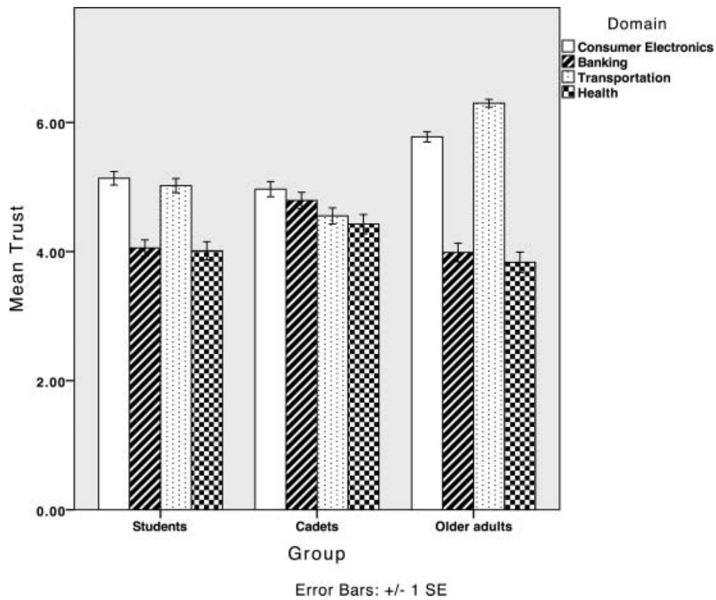


**Figure 1.** Trust as a function of group and automation domain.

automation was associated with higher trust ($M$ = 5.1, SD = 1.8) compared to decision automation ($M$ = 4.3, SD = 1.9). However, these main effects were qualified by several significant two-way interactions and one three-way interaction. All subsequent follow-up pairwise comparisons used to decompose interactions were adjusted for multiple comparisons (Sidak corrections) and used an alpha level of $p < .001$ as a criterion for significance.

### *Domain × user group interaction*

First, there was a significant two-way interaction of domain and group on trust (Figure 1). Follow-up pairwise comparisons of trust by domain within each group showed that students had significantly higher trust for both consumer electronics ($M$ = 5.1, SD = 1.6) and transportation ($M$ = 5.0, SD = 1.7) compared to banking ($M$ = 4.1, SD = 1.9) and health ($M$ = 4.0, SD = 2.2). The differences between consumer electronics and transportation and the differences between banking and health were not significantly different in students. Cadets showed no significant differences in trust by domain. Older adults reported significantly higher trust for both consumer electronics ($M$ = 5.8, SD = 1.0) and transportation ($M$ = 6.3, SD = 0.7) and lower trust for banking ($M$ = 4.0, SD = 1.7) and health ($M$ = 3.8, SD = 1.9). The differences between consumer electronics and transportation and the differences between banking and health were not significantly different in older adults.

Pairwise comparisons of trust by user group within each domain showed that within consumer electronics, older adults trust ($M$ = 5.5, SD = 1.0) was significantly higher than students ($M$ = 5.1, SD = 1.6) and cadets ($M$ = 5.0, SD = 1.7) while there was no difference between the cadets and students. In the banking domain, cadets had the highest level of trust ($M$ = 4.8, $SD$ = 1.8) compared to older adults ($M$ = 4.0, SD = 1.7) and younger adults ($M$ = 4.1, SD = 1.9) while students and older adults did not significantly differ from each other. Within the transportation domain, all group comparisons were significantly different with the older adults having the highest level of trust ($M = 6.3$, SD = 0.7), the younger adults at a moderate level ($M = 5.0$, SD = 1.7), and the cadets at the lowest level ($M = 4.6$, SD = 1.8). Finally, in the health domain, cadets ($M = 4.4$, SD = 1.9) trusted automation significantly more than older adults ($M = 3.8$, SD = 1.9) but there was no difference between cadets and students.

To summarise, student's and older adults' trust differed by domain with the highest levels of trust reserved for automation in consumer electronics and transportation compared to health and banking. Within consumer electronics and transportation, older adults had higher levels of trust than younger adults and for older adults, transportation was more trusted than consumer electronics while for younger adults, there was no significant difference. Interestingly, cadets showed no differentiation of trust by domain.

Students' and older adults' high trust with consumer technology was consistent with the finding that familiarity relates to trust (Muir and Moray 1996). High trust in the consumer electronics domain may have reflected the ubiquity of that class of technologies (Olson et al. 2011). However, the high trust in consumer electronics technology may have also reflected the lack of severe consequences for failure in the consumer electronics domain.

Student's and older adults' high trust of transportation automation was counterintuitive using the familiarity explanation. The automation examples in our scenarios were fairly modern and relatable but not expected to be widespread (e.g. blind spot alerts). The relative novelty of transportation technology compared to other domains was also shown in Olson et al. (2011). One explanation might be that due to the relative novelty for both civilian younger and older adults, users may have incomplete or inaccurate models of the capabilities and limitations of transportation automation (e.g. how it works). This inadequate mental model could lead to trust mis-calibration or over-trust. This would be consistent with studies that showed that mental model accuracy and quality affect automation misuse errors indicative of over-reliance and over-trust in both younger adults (Wilkinson, Fisk, and Rogers 2007) and older adults (Olson, Fisk, and Rogers 2009). Over trust in this situation is most accurately termed what Sheridan (2002) called *naive trust*. However, it was interesting that cadets did not show such differentiation by novelty. Their discipline and possibly access and exposure to more advanced defence-related automation in their training may have made them less prone to basing their trust on unfamiliarity or novelty.

Younger and older adults' low trust with health automation, however, was consistent with prior research that showed that when the cost of errors was high, as in health automation more than the other domains, users tended to decrease their reliance (Ezer, Fisk, and Rogers 2008) and trust in automation. In addition, older adults, who are more frequent users of health automation (Olson et al. 2011) might have been more familiar with the limitations, and may have adjusted their trust downward. Meanwhile, younger adults' low trust of health automation may have simply reflected unfamiliarity with the technology.

Finally, another explanation for low trust in health and banking automation for younger and older adults might have been that when the automation failed at such a seemingly routine but critical task in the health or banking domain (monitoring a health parameter, transfer of funds), user trust in the entire domain may have been irreparably damaged

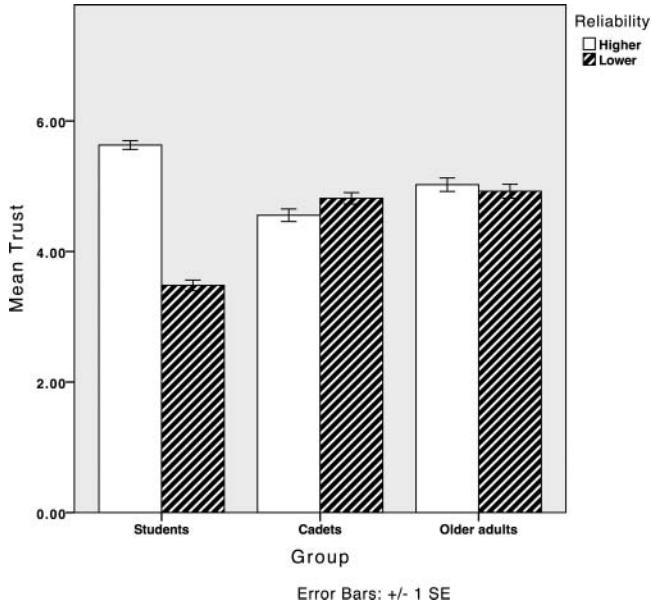


**Figure 2.** Trust as a function of group and reliability.

(Madhavan, Wiegmann, and Lacson 2006). This contrasts with the task requirements of automation in the other domains (consumer electronics and transportation) which required highly remote sensing in a fuzzy situation (weather radio) or relied on extremely fast reaction time (transportation); tasks which might be classified as more 'difficult' for humans as compared to machines. However, since we did not assess perceived difficulty of the task, this is speculation.

### *Automation reliability × user group interaction*

We hypothesised a relationship between reliability and trust, but not differentiated by group. However, the two-way interaction of automation reliability and group on trust showed a large effect (Figure 2). Follow-up pairwise comparisons showed that students had higher trust with more reliable automation ($M$ = 5.6, SD = 1.4) than less reliable automation ($M$ = 3.5, $SD$ = 1.8). Cadets and older adults showed no significant differences in trust as a function of automation reliability. Group comparisons showed that with reliable automation, students' trust ($M$ = 5.6, SD = 1.4) was significantly higher than older adults' trust ($M$ = 5.0, SD = 1.8) which in turn was significantly higher than cadets' trust ($M$ = 4.5, SD = 2.0). However, when presented with examples of unreliable automation, students' trust was significantly lower ($M$ = 3.5, SD = 1.8) compared to cadets ($M$ = 4.8, SD = 1.8) and older adults ($M$ = 4.9, SD = 1.9). There were no significant differences in trust of unreliable automation between cadets and older adults.

To summarise, students' trust in automation was sensitive to its reliability in the expected direction: more reliable automation was more trusted and less reliable automation was less trusted, consistent with the general finding in the literature on trust and reliability (e.g. Lee and Moray 1992; Hancock et al. 2011). However, cadets and older adults trust seemed insensitive to automation reliability; specifically, their trust remained relatively high when automation reliability was perceived to be lower. Rovira, McGarry, and Parasuraman (2007) also found trust insensitivity to reliability in their study with civilian students in a military task context.

Greater trust in automation than is warranted, or over trust, could be considered an indicator of complacency (Moray and Inagaki 2000; Parasuraman and Manzey 2010; Parasuraman, Sheridan, and Wickens 2000; Parasuraman, Sheridan, and Wickens 2008). Our finding of relatively higher trust ratings in older adults compared to younger adults was unsurprising and consistent with the literature on aging and automation (Ho, Wheatley, and Scialfa 2005; Ho, Kiff, Plocher, and Haigh 2005). Typically, the explanation for older adults' increased complacency with automation has been their reduced capacity for monitoring automation failures due to age-related declines in cognitive abilities such as working memory (Ho, Wheatley, & Scialfa). However, the lack of sensitivity in reliability among cadets, despite their youth, suggests that complacency effects in older adults may be attributable to additional factors other than cognitive aging. Cadets' lack of sensitivity to reliability may instead be a product of cadets treating automation as a teammate (Kennedy, Sibley, and Coyne 2015) and also due to the influence of organisational culture that mandates that authority and teammates be trusted.

The inability or unwillingness to adjust trust ratings to reliability among cadets seems, on the surface, inconsistent with a study by Wang, Jamieson, and Hollands (2009) who found that in the task context of combat identification (military domain), providing system reliability information led to appropriate shifts in trust scores (higher system reliability led to higher trust scores). But there were two crucial differences that make direct comparison difficult. First, the domain of the automated system in their study was military while ours was consumer-oriented automation (collapsed across all four domains). Automation in the military domain may have higher consequences for misuse or disuse compared to the consumer-oriented technology presented in our study. However, two of the domains used in our study (health and transportation) were of relatively higher severity in that the consequences of complacency should have been relatively higher than the other domains. Second, Wang, Jamieson, and Holland's participants who showed proper trust calibration were college students. In our study, college students showed proper trust calibration but our cadets did not distinguish between high and low reliability. For these reasons, the results from the students in our study are a more comparable reference group with the results of Wang, Jamieson, and Holland. That comparison showed that both groups' trust scores were sensitive to system reliability. Our result should place important limits on the generalisability of research carried out using college samples but intended for other groups (e.g. soldiers) as our students and cadets perceived automation differently based on reliability.

### *Automation stage × user group interaction*

The two-way interaction of automation stage and group was significant, illustrated in Figure 3, confirming our hypothesis 5. This two-way interaction was further qualified by domain (explained in next section) but because we hypothesised this specific effect, we will discuss it despite the presence of the three-way interaction. Pairwise comparisons showed that students' and older adults' trust in automation did not significantly vary by

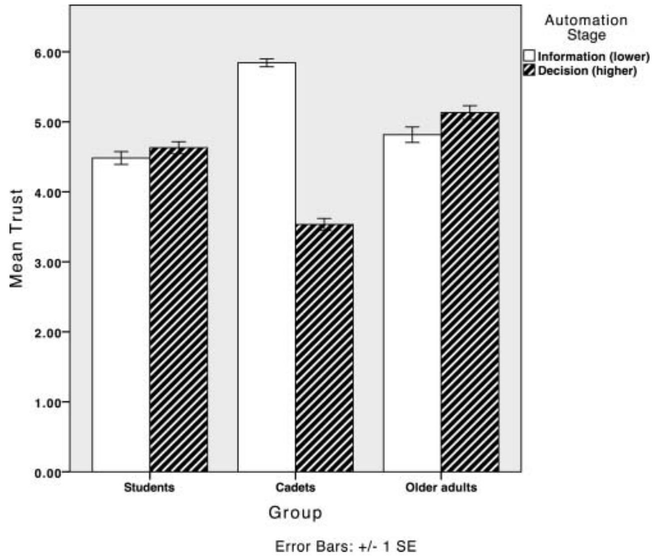


**Figure 3.** Trust as a function of group and automation stage.

stage of automation. However, cadets had significantly higher trust for information automation ($M$ = 5.8, SD = 1.2) than decision automation ($M$ = 3.5, SD = 1.8). Additional comparisons showed that within information automation, cadets trust ($M$ = 5.8, SD = 1.2) was significantly higher than students ($M$ = 4.5, SD = 2.0) and older adults ($M$ = 5.0, SD = 1.7). With decision automation, cadets' trust was significantly lower ($M$ = 3.5, SD = 1.8) than students ($M$ = 4.6, SD = 1.9) and older adults ($M$ = 5.1, SD = 1.7).

Cadets' significantly higher trust of information automation compared to students partially supported hypothesis 3 that due to training and organisational culture, cadets would have higher trust than students. These results were in opposition to hypothesis 5 that decision automation would be associated with more trust (Rovira, McGarry, and Parasuraman 2007); however this stage/trust relationship being limited to cadets was conceptually consistent with Parasuraman and Wickens' (2008) statement that trust may be lower with decision automation because they were referring to experienced users. Cadets' trust sensitivity to stage of automation could be attributed to organisational culture because civilian students and older adults do not show the same pattern (eliminating age and cohort as explanations). However, it was interesting that Rovira, McGarry, and Parasuraman found the stage–trust relationship in their study with civilian students in a military task domain. Decision automation tends to be more opaque and abstract because it replaces higher level cognitive processing such as decision-making. The opaqueness comes from the need for complex computational algorithms or more sensor fusion required with higher stages of automation. Cadets may be especially prone to distrust decision automation because their prior exposure with automation may include combat training systems where the severity or consequences are grave compared to civilians. For comparison, the three-way interaction (presented below) suggests that severity does indeed matter with distrust of decision automation being greatest in the high severity health domain.

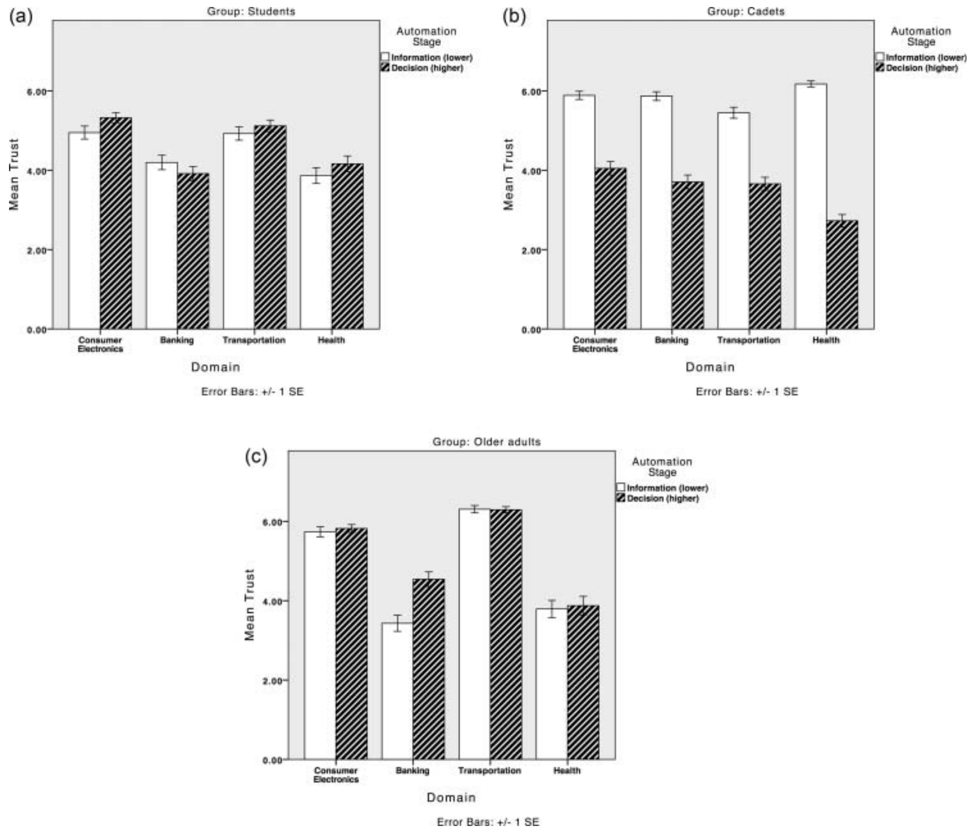


**Figure 4.** Trust as a function of domain, automation stage for (a) students, (b) cadets, and (c) older adults.

### *Group × domain × automation stage interaction*

The 3-way interaction of group × domain × automation stage was significant (Figure 4). The source of the interaction was a significant 2-way interaction between domain and automation stage within the cadets, $F(3,150) = 18.79$, $\eta_p^{\,2} = 0.27$, (Figure 4(a)), and older adults $F(3,105) = 5.90$, $\eta_p^{\,2} = 0.14$, (Figure 4(b)). The two-way interaction in students was marginally significant ($p = .006$) and thus was ignored (Figure 4(c)). Pairwise comparisons showed that for the cadets, the source of the significant two-way interaction between domain and stage was that with information automation, trust in health automation ($M = 6.2$, SD $= 0.8$) was only significantly *higher* than transportation ($M = 5.4$, SD $= 1.4$). But with decision automation, trust in health automation ($M = 2.7$, SD $= 1.6$) was significantly *lower* than all other domains. For older adults, the source of the significant two-way interaction between domain and stage was that with information automation, transportation ($M = 6.3$, SD $= 0.8$) was trusted significantly higher than banking ($M = 4.5$, SD $= 1.6$) and health ($M = 3.8$, SD $= 1.8$). But with decision automation, transportation ($M = 6.3$, SD $= 0.7$) was trusted significantly more than all other domains. In contrast to cadets, who showed a strong stage effect on trust, older adults only showed a reverse stage effect in the banking domain.

To summarise, students' trust in automation was insensitive to the stage of automation. However, cadets and older adults' trust in automation appears to be different and domain and automation stage are critical moderators. Cadets showed a strong overall effect of stage such that trust was lower with decision automation with all domains of automation but especially health automation showing the largest difference by stage. Recall that the health domain was thought to be most severe and have the worse consequences for failure.

Older adults' trust, however, was relatively insensitive to stage except for the banking domain where, contrary to cadets, decision automation was more trusted than information automation. In addition, they showed a high level of trust in the transportation domain, regardless of the stages. Older adults' high trust in the transportation domain was consistent with a study by Donmez, Boyle, Lee, and McGehee (2006) that showed older drivers are trusting of driver distraction mitigation automation even when it was shown to be imperfect. The authors surmised that older adults, due to age-related cognitive impairments, welcomed any driving-related safety automation because it allowed them to maintain their independence, no matter the reliability of the automation. This may also explain older adults' higher trust of decision automation in banking; managing finances can be a cognitively demanding task, older adults may be more willing to offload more of the task to automation, despite occasional reliability issues, because it means independence. The alternative to maintaining a high level of trust in both domains would be to lower trust, disuse automation, and potentially lose control of their finances or mobility. Driving is the most frequent mode of transport for those above age 65 (Rosenbloom and Waldorf 2001). It is also a major component of older adults' sense of functional independence (Delligner et al. 2001) so the prospects of losing control over mobility, and perhaps their finances, would be highly undesirable for older adults.

In summary, the results paint a complex picture of the effect of automation domain of varying reliability and stages on trust in different groups of users. First, students' and older adults' trust differed by domain with the highest levels of trust reserved for the consumer electronics and transportation domains. Second, it appeared that students' trust, but not cadets' or older adults', was sensitive to automation reliability. Finally, cadets, and older adults in one domain, were particularly sensitive to stage of automation in opposite ways.

## Conclusion

The results of the study showed that trust in automation was highly dependent on the interaction of domain, stage, and reliability of automation but also the user group. Students were able to successfully calibrate their trust ratings depending on the reliability of the system while older adults and cadets were not. It was unsurprising that aging adults were unable to calibrate their trust appropriately (Ho, Wheatley, and Scialfa 2005), but the finding that cadets were unable to calibrate trust appropriately may be attributable to the cultural differences that come with being in a military organisation. While we cannot be exactly sure what those differences are, prior research does suggest some of the ways in which military populations can differ from civilian (Kennedy, Sibley, and Coyne 2015; Hollands and Neyedli 2011; Soeters, Winslow, and Weibull 2006). Further, we could attribute differences to organisational culture based on the comparisons with students, who were of the same age and cohort yet were able to distinguish between low and high reliability. We could also rule out any pre-existing individual differences in complacency potential as there were no significant differences in the measure of complacency potential between cadets and students.

The differential effect of stages of automation on trust was particularly interesting; cadets showed dramatically lowered trust for decision automation across all the domains presented in the study. Civilian students showed no such effect of stage of automation on trust. This unique effect of organisational culture on trust by stage of automation places important limits on our ability to generalise the results of automation trust studies using civilians. Similarly, older adults showed an extremely targeted difference in trust by automation stage in the banking domain with higher trust for decision automation. Another interesting finding was that although consumer electronics, transportation, and health did not show any differences by stage, older adults had an extremely high level of trust in transportation automation. This high trust might have been due to the uniquely older adult issue of the prospect of losing independence.

Despite the caution suggested by these findings, it must be put into context with the limitations of this study. First, we used a scenario methodology where participants read examples of a hypothetical user interacting with automation in contrast to actually experiencing several trials of automation. Thus, there was some danger of demand characteristics; that is, participants may have detected the manipulation and adjusted their trust responses to what they think we expected of them (i.e. higher trust with higher reliability). However, we included a question at the end of the study that asked participants to report the purpose of the study and none were precise. Instead, they had generic responses such as, ‘trust in technology’, or ‘comfort with the use of technology’, Finally, the use of scenarios is a well-accepted methodology when the goal is to assess subjective, evaluative judgments and perceptions of complex and multidimensional phenomenon (Rossi and Anderson 1982; Jasso 2006; Atzmüller and Steiner 2010), and has been used in prior studies to examine the topic of trust and other perceptions of automation (Endsley and Kiris 1995; Mosier and Fischer 2012; Pak, McLaughlin, and Bass 2014).

Another possible issue with this study was that we frequently appealed to inferences about personal characteristics about each group to explain observed effects, although we did not directly measure those characteristics (e.g. the discipline of military cadets, the fear of losing independence in older adults). However, the characteristics described were supported in the literature by other studies of those groups (cadets, older adults). This may place some constraints on the validity of the explanations for the observed effects but provides guidance for future replication.

Finally, the domains used in this study are differentially regulated by the government and trust in the technologies may have reflected trust in those institutions but also perceptions of the government’s ability to regulate them. This would clearly affect one’s predisposition to trust automation (Lee and See 2004). For example, in the United States, three of the four domains used in this study (transportation, health, and banking) are highly regulated by one or more government agencies devoted to them. However, the results did not neatly differentiate highly regulated versus less-regulated domains (e.g. consumer electronics). For example, civilians’ low trust in the highly regulated domains of banking and health was contrasted by their high trust in transportation. A future study might disambiguate these sources of trust perceptions from the technology itself.

One curious finding was that there were no age differences in the CPRS. We expected to find higher CPRS scores for older adults, both reflecting their generally increased tendency to be complacent but also replicating other studies that have found age differences on that measure (e.g. Pak, McLaughlin, and Bass 2014). In Pak, McLaughlin, and Bass' study, both younger and older adults CPRS scores ($M = 43.4$, SD = 4.6 and $M = 47.3$, SD = 4.7, respectively) were lower than the values reported here ($M = 51.5$, SD = 3.6 and $M = 50.6$, $SD = 5.3$, respectively). The younger and older adults were obviously different in both studies but the students came from the same University with the same mean age (18) and the older adults came from the same geographic area (south-eastern USA) again with about the same mean age between studies (early 70s). The other major difference was the passage of about three years between the two studies (2012 versus 2015). This suggests the interesting possibility that the general public is becoming more complacent as automation becomes more frequently encountered in daily life. Even more, that younger adults are becoming as complacent as older adults. Regardless of the age-equivalence on CPRS found in this study, we still did find evidence of age differences in trust (an indicator of complacency), such that older adults had higher trust than the students. An interesting future study might document the long-term changes in trust and attitudes toward automation and its relation to frequency of technology usage and exposure.

The findings highlight the importance of representative design, especially in human factors and ergonomics research. The principle of representative design applied here means that in the design of trust and automation experiments, it is crucial to accurately sample a wide range users' characteristics under a wide variety of conditions, or domains. Representative design is fundamental to the ability to generalise results from singular studies to other situations, domains, and groups (Hammond 1998; Fisk and Kirlik 1996; Czaja and Sharit 2003). The results also suggest some caution in generalising extant research into other domains or users.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Funding

This research was supported by a grant from the Air Force Office of Scientific Research [award number FA9550-12-1-0385].

## About the authors

*Richard Pak* is an associate professor in the Department of Psychology at Clemson University. He received his Ph.D. in psychology in 2005 from the Georgia Institute of Technology.

*Ericka Rovira* is an associate professor in the Department of Behavioral Sciences & Leadership at the US Military Academy, West Point. She received her Ph.D. in applied experimental psychology from the Catholic University of America in 2006.

*Anne McLaughlin* is an associate professor in the Department of Psychology at North Carolina State University. She received her Ph.D. in psychology in 2007 from the Georgia Institute of Technology.

*Natalee Baldwin* is a Ph.D. student in the Department of Psychology at Clemson University. She received her MS in human factors in 2014 from Clemson University.

## ORCID

*Richard Pak* http://orcid.org/0000-0001-9145-6991

## References

Atzmüller, C., and P.M. Steiner. 2010. "Experimental Vignette Studies in Survey Research." *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences* 6 (3): 128–138. http://doi.org/10.1027/1614-2241/a000014

Calhoun, G.L., M.H. Draper, and H.A. Ruff. 2009. "Effect of Level of Automation on Unmanned Aerial Vehicle Routing Task." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 53 (4): 197–201. http://doi.org/10.1177/154193120905300408

Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). New York: Academic Press.

Crocoll, W.M., and B.G. Coury. 1990. "Status or Recommendation: Selecting the Type of Information for Decision Aiding." In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 1524–1528. Santa Monica.

Czaja, S.J., and J. Sharit. 2003. "Practically Relevant Research: Capturing Real World Tasks, Environments, and Outcomes." *The Gerontologist* 43 (1): 9–18.

Dellinger, A.M., M. Sehgal, D.A. Sleet, and E. Barrett Connor. 2001. "Driving Cessation: What Older Former Drivers Tell us." *Journal of the American Geriatrics Society* 49 (4): 431–435. http://doi.org/10.1046/j.1532-5415.2001.49087.x

Donmez, B., L.N. Boyle, J.D. Lee, and D.V. McGehee. 2006. "Drivers' Attitudes Toward Imperfect Distraction Mitigation Strategies." *Transportation Research part F: Traffic Psychology and Behaviour* 9 (6): 387–398.

Dzindolet, M., S.A. Peterson, R.A. Pomranky, L.G. Pierce, and H.P. Beck. 2003. "The Role of Trust in Automation Reliance." *International Journal of Human-Computer Studies* 58 (6): 697–718. http://doi.org/10.1016/S1071-5819(03)00038-7

Dzindolet, M., L.G. Pierce, H.P. Beck, L.A. Dawe, and B.W. Anderson. 2001. "Predicting Misuse and Disuse of Combat Identification Systems." *Military Psychology* 13 (3): 147–164. http://doi.org/10.1207/S15327876MP1303_2

Endsley, M.R., and E.O. Kiris. 1995. "The Out-of-the-Loop Performance Problem and Level of Control in Automation." *Human Factors* 37 (2): 381–394. http://doi.org/10.1518/001872095779064555

Ezer, N., A.D. Fisk, and W.A. Rogers. 2008. "Age-Related Differences in Reliance Behavior Attributable to Costs Within a Human-Decision Aid System." *Human Factors* 50 (6): 853–863. http://doi.org/10.1518/001872008X375018

Faul, F., and E. Erdfelder. 1992. GPOWER: *A Priori, Post-hoc, and Compromise Power Analyses for MS-DOS [Computer program]*. Bonn: Department of Psychology, Bonn University.

Fisk, A.D., and A. Kirlik. 1996. "Practical Relevance and Age-Related Research: Can Theory Advance Without Practice?" In *Aging and skilled performance: Advances in theory and application*, edited by W.A. Rogers, A.D. Fisk, and N. Walker, 1–15. Mahwah, NJ: Erlbaum.

Fitts, P., ed. 1951. *Human Engineering for an Effective Air Navigation and Traffic Control System*. Washington, DC: National Research Council.

Hammond, K.R. 1998. “Representative Design.” *Retrieved October* 9: 1998. http://brunswik.org/notes/essay3.html

Hancock, P.A., D.R. Billings, K.E. Schaefer, J.Y.C. Chen, E.J. de Visser, and R. Parasuraman. 2011. “A meta-analysis of factors affecting trust in human-robot interaction.” *Human Factors* 53 (5): 517−527. http://doi.org/10.1177/0018720811417254

Ho, G., L.M. Kiff, T. Plocher, and K.Z. Haigh. 2005. “A Model of Trust & Reliance of Automation Technology for Older Users.” In *AAAI-2005 Fall Symposium:Caring Machines: AI in Eldercare*, 45−50. Washington, DC.

Ho, G., D. Wheatley, and C.T. Scialfa. 2005. “Age Differences in Trust and Reliance of a Medication Management System.” *Interacting with Computers* 17 (6): 690−710. http://doi.org/10.1016/j.intcom.2005.09.007

Hoff, K.A., and M. Bashir. 2015. “Trust in Automation Integrating Empirical Evidence on Factors That Influence Trust.” *Human Factors* 57 (3): 407−434. http://doi.org/10.1177/0018720814547570

Hofstede, G. 1991. *Culture’s Consequences: Software of the Mind*. London: McGrawHill.

Hollands, J.G., and H. Neyedli. 2011. “A Reliance Model for Automated Combat Identification Systems: Implications for Trust in Automation.” In *Trust in Military Teams,* edited by N.A. Stanton, 151−182. Surrey: Ashgate.

Inagaki, T. 2006. “Design of Human−Machine Interactions in Light of Domain-Dependence of Human-Centered Automation.” *Cognition, Technology & Work* 8 (3): 161−167. http://doi.org/10.1007/s10111-006-0034-z

Jasso, G. 2006. “Factorial Survey Methods for Studying Beliefs and Judgments.” *Sociological Methods & Research* 34 (3): 334−423.

Kennedy, W.G., C. Sibley, and J. Coyne. 2015. “Self-Confidence of Autonomous Systems in a Military Environment.” *2015 AAAI Fall Symposium Series*. Arlington, VA.

Kincaid, J.P., R.P. Fishburne Jr., R.L. Rogers, and B.S. Chissom. 1975. *Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel*. No. RBR-8-75. Millington TN: Naval Technical Training Command, Research Branch.

Lee, J., and N. Moray. 1992. “Trust, Control Strategies and Allocation of Function in Human-Machine Systems.” *Ergonomics* 35 (10): 1243−1270.

Lee, J.D., and N. Moray. 1994. “Trust, Self-Confidence, and Operators’ Adaptation to Automation.” *International Journal of Human-Computer Studies* 40 (1): 153−184.

Lee, J.D., and K.A. See. 2004. “Trust in Automation: Designing for Appropriate Reliance.” *Human Factors: The Journal of the Human Factors and Ergonomics Society* 46 (1): 50−80.

Madhavan, P., D.A. Wiegmann, and F.C. Lacson. 2006. “Automation Failures on Tasks Easily Performed by Operators Undermine Trust in Automated Aids.” *Human Factors* 48 (2): 241−256.

McBride, S.E., W.A. Rogers, and A.D. Fisk. 2011. “Understanding the Effect of Workload on Automation use for Younger and Older Adults.” *Human Factors* 53 (6): 672−686.

Melenhorst, A.-S.S., W.A. Rogers, and D.G. Bouwhuis. 2006. “Older Adults’ Motivated Choice for Technological Innovation: Evidence for Benefit-Driven Selectivity.” *Psychology and Aging* 21 (1): 190−195. http://doi.org/10.1037/0882-7974.21.1.190

Moray, N., and T. Inagaki. 2000. “Attention and Complacency.” *Theoretical Issues in Ergonomics Science* 1 (4): 354−365.

Mosier, K. and U. Fischer. 2012. Impact of Automation, Task and Context Features on Pilots’ Perception of Human-Automation Interaction.” *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 56 (1): 70−74. http://doi.org/10.1177/1071181312561035

Muir, B.M., and N. Moray. 1996. “Trust in Automation. Part II. Experimental Studies of Trust and Human Intervention in a Process Control Simulation.” *Ergonomics* 39 (3): 429−460. http://doi.org/10.1080/00140139608964474

Olson, K.E., A.D. Fisk, and W.A. Rogers. 2009. “Collaborative Automated Systems: Older Adults’ Mental Model Acquisition and Trust in Automation.” *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 53 (22): 1704−1708 http://doi.org/10.1177/154193120905302210

Olson, K.E., M.A. O'Brien, W.A. Rogers, and N. Charness. 2011. "Diffusion of Technology: Frequency of use for Younger and Older Adults." *Ageing international* 36 (1): 123–145. http://doi.org/10.1007/s12126-010-9077-9

Onnasch, L., C.D. Wickens, H. Li, and D. Manzey. 2013. "Human Performance Consequences of Stages and Levels of Automation an Integrated Meta-Analysis." *The Journal of the Human Factors and Ergonomics Society* 56 (3): 476–488. http://doi.org/10.1177/0018720813501549

Pak, R., N. Fink, M. Price, B. Bass, and L. Sturre. 2012. "Decision Support Aids with Anthropomorphic Characteristics Influence Trust and Performance in Younger and Older Adults." *Ergonomics* 55 (9): 1059–1072. http://doi.org/10.1080/00140139.2012.691554

Pak, R., A.C. McLaughlin, and B. Bass. 2014. "A Multi-Level Analysis of the Effects of Age and Gender Stereotypes on Trust in Anthropomorphic Technology by Younger and Older Adults." *Ergonomics* 57 (9): 1277–1289. http://doi.org/10.1080/00140139.2014.928750

Parasuraman, R., and D.H. Manzey. 2010. "Complacency and Bias in Human Use of Automation: An Attentional Integration." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 52 (3): 381–410.

Parasuraman, R., R. Molloy, and I.L. Singh. 1993. "Performance Consequences of Automation-Induced Complacency." *The International Journal of Aviation Psychology* 3 (1): 1–23.

Parasuraman, R., and V. Riley. 1997. "Humans and Automation: Use, Misuse, Disuse, Abuse." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 39 (2): 230–253.

Parasuraman, R., T.B. Sheridan, and C.D. Wickens. 2000. "A Model for Types and Levels of Human Interaction with Automation." *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans* 30 (3): 286–297. http://doi.org/10.1109/3468.844354

Parasuraman, R., T.B. Sheridan, and C.D. Wickens. 2008. "Situation Awareness, Mental Workload, and Trust in Automation: Viable, Empirically Supported Cognitive Engineering Constructs." *Journal of Cognitive Engineering and Decision Making* 2 (2): 140–160. http://doi.org/10.1518/155534308X284417

Parasuraman, R., and C.D. Wickens. 2008. "Humans: Still vital after all these Years of Automation." *Human Factors* 50 (3): 511–520. http://doi.org/10.1518/001872008X312198

Petrinovich, L. 1989. "Representative Design and the Quality of Generalization." In *Everyday Cognition in Adulthood and Late Life*, edited by L.W. Poon, D.C. Rubin and B.A. Wilson, 11–24. Cambridge: Cambridge University Press.

Richardson, J. 2011. "Eta Squared and Partial eta Squared as Measures of Effect Size in Educational Research." *Educational Research Review* 6: 135–147.

Rogers, W.A., R. Pak, and A.D. Fisk. 2007. "Applied Cognitive Psychology in the Context of Everyday Living." In *Handbook of Applied Cognition*, edited by F. Durso, 1–27. Hoboken, NJ: John Wiley & Sons.

Rosenbloom, S., and B. Waldorf. 2001. "Older Travelers: Does Place or Race Make a Difference?" Transportation Research Circular (Number E-C026). Washington, DC: Transportation Research Board.

Rossi, P.H., and A.B. Anderson. 1982. "The Factorial Survey Approach: An Introduction." *Measuring social judgments: The factorial survey approach*. 15–67.

Rovira, E., K. McGarry, and R. Parasuraman. 2007. "Effects of Imperfect Automation on Decision Making in a Simulated Command and Control Task." *Human Factors* 49 (1): 76–87.

Sarter, N.B., and B. Schroeder. 2001. "Supporting Decision Making and Action Selection Under Time Pressure and Uncertainty: The case of in-Flight Icing." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 43 (4): 573–583.

Sauer, J., and B. Rüttinger. 2007. "Automation and Decision Support in Interactive Consumer Products." *Ergonomics* 50 (6): 902–919. http://doi.org/10.1080/00140130701254266

Sheridan, T.B. 2002. *Humans and Automation: System Design and Research Issues*. Santa Monica, CA: John Wiley & Sons.

Sheridan, T.B., and W.L. Verplank. 1978. *Human and Computer Control of Undersea Teleoperators*. Cambridge: Massachusetts Institute of Technology, Man Machine Systems Laboratory.

Singh, I.L., R. Molloy, and R. Parasuraman. 1993. "Individual Differences in Monitoring Failures of Automation." *The Journal of General Psychology* 120 (3): 357–373.

Soeters, J.L., D.J. Winslow, and A. Weibull. 2006. "Military Culture." In *Handbook of Sociology and Social Research*, edited by G. Caforio, 237–254. Boston, MA: Springer. http://doi.org/10.1007/0-387-34576-0_14

Wang, L., G.A. Jamieson, and J.G. Hollands. 2009. "Trust and reliance on an automated combat identification system: The role of aid reliability and reliability disclosure." *Human Factors* 51 (3): 281–291.

Warren, N.D. 1956. "Automation, Human Engineering, and Psychology." *American Psychologist* 11 (10): 531–536.

Wickens, C.D., H. Li, A. Santamaria, A. Sebok, and N.B. Sarter. 2010. "Stages and Levels of Automation: An Integrated Meta-analysis." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 54 (4): 389–393. http://doi.org/10.1177/154193121005400425

Wilkison, B.D., A.D. Fisk, and W.A. Rogers. 2007. "Effects of Mental Model Quality on Collaborative System Performance." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 51 (22): 1506–1510. http://doi.org/10.1177/154193120705102208