Ergonomics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/terg20

Enhancing component-specific trust with consumer automated systems through humanness design

Jeremy Lopez, Heather Watkins & Richard Pak

ERGONOMICS
https://doi.org/10.1080/00140139.2022.2079728

Taylor & Francis
Taylor & Francis Group

Check for updates

ARTICLE

Enhancing component-specific trust with consumer automated systems through humanness design

Jeremy Lopez, Heather Watkins and Richard Pak

Department of Psychology, Clemson University, Clemson, SC, USA

ABSTRACT
Consumer automation is a suitable venue for studying the efficacy of untested humanness design methods for promoting specific trust in multi-component systems. Subjective (trust, self-confidence) and behavioural (use, manual override) measures were recorded as 82 participants interacted with a four-component automation-bearing system in a simulated smart home task for two experimental blocks. During the first block all components were perfectly reliable (100%). During the second block, one component became unreliable (60%). Participants interacted with a system containing either a single or four simulated voice assistants. In the single-assistant condition, the unreliable component resulted in trust changes for every component. In the four-assistant condition, trust decreased for only the unreliable component. Across agent-number conditions, use decreased between blocks for only the unreliable component. Self-confidence and overrides exhibited ceiling and floor effects, respectively. Our findings provide the first evidence of effectively using humanness design to enhance component-specific trust in consumer systems.

Practitioner summary: Participants interacted with simulated smart-home multi-component systems that contained one or four voiced assistants. In the single-voice condition, one component's decreasing reliability coincided with trust changes for all components. In the four-voice condition, trust decreased for only the decreasingly reliable component. The number of voices did not influence use strategies.

Abbreviations: ACC: adaptive cruise control; CST: component-specific trust; SWT: system-wide trust; UAV: unmanned aerial vehicle; CPRS: complacency potential rating scale; MANOVA: multivariate analysis of variance

ARTICLE HISTORY
Received 29 November 2021
Accepted 6 May 2022

KEYWORDS
Human-machine systems; human-automation interaction; trust in automation; humanness

Introduction

Systems containing automation are not only increasingly present, but more complex, oftentimes consisting of several automated subsystems. For example, automobiles of the past would feature a single automated system such as adaptive cruise control (ACC). Current commercially available vehicles are likely to feature multiple independently automated subsystems, including collision warning systems, ACC, and lane-keeping assistance. Whereas the older vehicle can be considered a single-component system (for it contains a single automated subsystem), the current vehicle with multiple automated subsystems can be considered a multi-component system. Importantly, the complexity of a system's software does not necessarily determine whether a user will consider a system to be single- or multi-component; a single function like ACC incorporates any number of underlying algorithms. Instead, the apparent complexity of a system helps a user to develop a mental model of the system (Endsley, Bolte, and Jones 2003; Endsley 2017). Mental models are knowledge structures of the 'causal interconnections involving actions and environmental events that influence the functioning of the system' (Durso and Gronlund 1999, 297–298). A system's apparent complexity is perceived via the system interface and can be learned by interacting with the system, which then helps to develop and refine a mental model of how the entire system can be partitioned into subsystems (i.e. components) (Moray 1998). It is crucial for users to develop an accurate mental model of a system because a mental model that is consistent with a system's functionality establishes a base for the development of trust and acceptance (Kazi et al. 2007; Itoh 2012; Beggiato and Krems 2013). We adopt Lee and See's (2004) definition of trust in automation as

CONTACT Richard Pak richpak@clemson.edu Department of Psychology, Clemson University, Clemson, SC 29634, USA

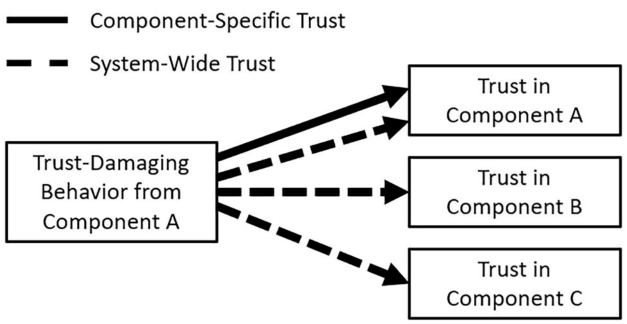

Figure 1. Diagram displaying the difference between component-specific trust (CST) and system-wide trust (SWT). With CST (indicated by the solid line), a trust-damaging behaviour from Component A influences trust in Component A. With SWT (indicated by the solid and dashed lines), a trust-damaging behaviour from Component A influences trust in every component.

'the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability' (54). Despite the wealth of literature on trust in human-automation interaction, additional research must explore how users trust in multi-component systems.

An unresolved question is whether multi-component system users discriminate between discrete and independent automated subsystems or treat the collection of automated subsystems as a singular system. That is, when one subsystem fails (e.g. ACC), do users discriminate and distinguish their trust in that component from their trust in other components? A user who loses trust in only the ACC is using a component-specific trust (CST) strategy, which suggests that their trust in any particular subsystem is unaffected by their trust in other subsystems. Alternatively, a user that loses trust in the ACC and other subsystems is using a system-wide trust (SWT) strategy, suggesting that one underperforming subsystem causes them to lose trust in other subsystems (Keller and Rice 2009; see Figure 1).

Another current limitation with the findings on trust strategies within multi-component systems (e.g. Keller and Rice 2009; Walliser, de Visser, and Shaw 2016; Rice et al. 2016; Foroughi et al. 2019) is the narrow focus on specific domains (typically industrial or military), which may limit the generalisability to more consumer-friendly applications. Consumer automation is any automated system manufactured to be used by an untrained user in a typically domestic environment, many of which can be considered multi-component systems (e.g. smart speakers, modern cellular phones). Compared to other domains, consumer automation is generally used for low-risk and highly interruptible environments and tasks (Sauer and Rüttinger 2007). Therefore, findings from studies of industrial and military automation may not always extend to the consumer domain. For example, designers are generally recommended to avoid highly automating systems related to decision making (Wickens et al. 1998). Sauer and Rüttinger (2007) found that a vacuum featuring highly automated decision automation standardised performance between motivated and unmotivated individuals, extending Wickens et al.'s (1998) suggestion to consumer automation. However, Pak et al. (2017) showed that, at least in single-component systems, trust can vary between domains, suggesting that domain can influence trust. Therefore, the findings from the military and industrial multi-component system literature may only be applicable to their respective domains. Users may decide to disuse an automated system they do not trust (Parasuraman and Riley 1997), consequently emphasising any factors that influence trust. A more complete understanding of trust in multi-component systems is only possible with the inclusion of systems beyond the typical domains.

Currently, there does not appear to be theoretical guidance as to when users tend to adopt a CST or SWT strategy. This limitation impedes application because the strategy choice may have performance implications for different kinds of systems. For example, Bean, Rice, and Keller (2011) sought to determine how subsystems with varying reliabilities influence performance on a flight simulation task. A user's trust in an automated system tends to decrease with decreasing system reliability (Lee and See 2004), where reliability is numerically represented as the percentage of events when the automation correctly and accurately performs an intended function. Estimates of the exact value vary, but trust significantly decreases once reliability decreases to a range between 70% (Kantowitz, Hanowski, and Kantowitz 1997) and 60% (Fox 1996). Bean, Rice, and Keller (2011) found that users' adoption of SWT led to worsened task performance with all subsystems regardless of reliability. This is especially important since, as previously mentioned, many consumer facing systems are now composed of multiple subsystems. The purpose of this research was to examine if users are more likely to adopt CST or SWT when interacting with a consumer-oriented multi-component system that incorporates elements of humanness design.

Do users engage in system-wide trust or component-specific trust?

Early studies of multi-component systems showed that users tended to adopt a CST strategy (Lee and Moray

1992, 1994). That is, trust in reliable subsystems was unaffected by the performance of an unreliable subsystem. In their studies, Lee and Moray examined trust in an automated process control scenario. Participants had to monitor a simulated pasteurisation plant with three subsystems that could be manually controlled or automated. One of the subsystems would occasionally fail to perform an action while the other two subsystems maintained perfect performance. One major finding is that changes in trust in one subsystem can occur independently of changes in trust for others (i.e. CST).

However, more recent studies with aviation gauge-monitoring tasks (Keller and Rice 2009; Rice and Geels 2010; Geels-Blair, Rice, and Schwark 2013) have shown that users tend to adopt a SWT strategy. That is, participants did not seem to completely discriminate between a single faulty automated gauge monitor from the others, leading to a decrement or ‘pull down’ of trust. This ‘pull down’ causes participants to lose trust in not only the unreliable subsystem, but also a lesser amount in the reliable subsystems. Subsequent studies have replicated this ‘pull down’ effect in unmanned aerial vehicle (UAV) assisted identification tasks (Walliser, de Visser, and Shaw 2016; Kluck et al. 2018; Foroughi et al. 2019) and transportation automation (Rice et al. 2016). Furthermore, the body of research suggests that with reliable subsystems, SWT can lead to increased response times, worsened task performance, and increased verification behaviours.

Given the potential consequences of SWT and the ‘pull down’ effect, research has centred on the factors that affect whether users adopt a CST or SWT strategy. Although SWT has primarily been considered a negative outcome, a system designer may want to use these factors to promote SWT. For example, a processing plant with functionally independent subsystems may require all subsystems to be functioning for safe operation; in such a situation, it may be preferable for an operator to cease use of all subsystems once one becomes unreliable. The difficulty with preventing SWT is that any amount of contamination between components leads to SWT. For example, Walliser, de Visser, and Shaw (2016) tested the effectiveness of providing performance feedback for every component as a method for promoting CST. They found that although the feedback prevented participants from equally trusting every component, the presence of one unreliable component lowered trust in reliable components. That is, participants appeared to be discriminating between components, but not discriminating enough to view each component as an independent and distinct agent. Therefore, it appears that the only way for CST to be achieved is for a user to completely discriminate between components. Later studies attempted to maximise component discrimination, but users still implemented SWT (Rice et al. 2016; Kluck et al. 2018). For instance, Kluck et al. (2018) had participants interact with four UAVs to complete a target identification task. Inspired by the Gestalt Principle of Organisation and related work (Campbell 2007), the researchers assigned half the participants to complete the task with four visually similar UAVs while the remainder interacted with four UAVs which differed in colour, wing-shape, and size. The researchers found that compliance and trust ratings indicated SWT in both conditions. One conclusion from this finding is that the visual dissimilarity of the UAVs was insufficient to promote component discriminability. The participants likely registered the visual uniqueness of each UAV, but still considered each one as just another UAV. Instead, methods for promoting discriminability may need to influence users to perceive components as fundamentally unique.

One potential way to enhance the discriminability of automation components might be to incorporate aspects of humanness design (de Visser et al. 2018). Humanness design encompasses any feature that is included with the purpose of connecting and communicating with the human user. For example, voiced smart assistants featured in consumer technologies (e.g. Google Assistant on Android devices, Apple’s Siri on iOS devices) contain two types of humanness design: simulated human speech and adaptation to user behaviours. Implementing humanness design into complex systems can reduce the likelihood of a discrepancy between user’s perceptions of a system and the system’s true capabilities (de Visser et al. 2018), but it may also encourage users to treat discrete components as separate elements. That is, humanness design may be capable of promoting users to accurately lose trust in only unreliable components (i.e. CST) instead of inaccurately losing trust in the entire system (i.e. SWT).

The current study

In the current study, we propose to use design concepts borrowed from the literature in humanness (de Visser et al., 2018) to enhance the discriminability of subsystems. By humanising subsystems and encouraging their perception as separate agents, we may maximally reinforce to the user that the systems are individual and unique, thereby encouraging a CST strategy. We humanised the automation in this study by using simulated smart assistants, each with unique names and voice characteristics. This kind of manipulation is already commonly used in many consumer domains (e.g. voice assistant, car interfaces) and so is a natural fit for studying multi-component systems in consumer domains. Prior research has shown that users can differentiate between multiple voices from an automated system, even if the voices are presented from the same physical device (Nass, Steuer, and Tauber 1994). Therefore, the implementation of multiple voiced agents into a system may provide enough discriminability for users to adopt CST.

In the current study, we presented participants with a simulated smart home system responsible for controlling lighting conditions in multiple rooms of a hypothetical house. Participants in one condition interacted with a single simulated voice assistant, whereas the remaining participants interacted with multiple simulated voice assistants (one assistant per room). We measured our primary variable, trust, alongside other subjective (self-confidence) and behavioural measures (e.g. use). Our prediction was that users who interact with multiple automated subsystems that are differentiable (via humanness cues) will tend to adopt a strategy closer to CST. More specifically, we hypothesised that the presence of an unreliable component would influence trust and related variables for participants interacting with the single-voiced system, but no such contamination would be present for users of the multi-voiced system.

Method

Participants

Eighty-seven undergraduate students at a southeastern university completed the study. Of the 87 participants, five participants' data were removed due to incomplete data. The average age of the remaining 82 students (43 females, 39 males) was 19.17 ($SD = 1.49$). Participants were given course credit for completing the study.

Materials

Data was collected in a lab setting on desktop computers. Participants interacted with the system using a mouse and keyboard. To ensure the audibility of the smart assistants, participants wore wired over-ear headphones during the duration of the experimental task.

Task

Prior work on multi-component systems placed the operator in a supervisory role, requiring the operator to monitor the automation as it performs its tasks. Therefore, one of our goals was to simulate a supervisory task within the consumer domain. The experiment used a low-fidelity smart home simulator designed for the current study. The task is designed to simulate the use of a smart home mobile application to adjust light intensities in various rooms of a house. The simulator (Figure 2) consisted of four primary areas: the phone interface (left), the task list (right), instructions (top center), and the 'NEXT' button to end the current trial (bottom right). A researcher verbally presented the task instructions before the experimental task began, but the instructions were included to remind the participants how to operate the application. The task list displayed the tasks that needed to be completed during a given trial. All tasks were listed in this room order: Master Bedroom, Living Room, Kitchen, and Dining Room. Each task list required the participant to adjust each room's light intensities to a random value between 0% and 100%.

Participants had the ability to either manually perform or automate tasks. If the participant chose to manually complete the task, they would click the 'Rooms' button on the phone interface, taking them to the 'Rooms' screen (Figure 3). This screen listed all four rooms' current lighting conditions, with sliders for adjusting lighting intensity. The 'Return' button would return the participant to the phone application's 'Home' screen. If the participant chose to automate a task, they would click the 'Assistance' button on the phone interface. From here the participants could decide which rooms' lightings to be automated.

Once a participant had completed all the tasks, they could end the trial by clicking the 'NEXT' button in the bottom right corner of the interface (Figure 2) to open the feedback window (Figure 4). If a task was manually completed, then the participant would receive text feedback about the success on the previous trial. If a task was automated, then vocal feedback from a smart assistant would inform the participant about success on the previous trial.

Automation perceptions and trust measures

Participants completed the Complacency Potential Rating Scale (CPRS; Singh, Molloy, and Parasuraman 1993) to indicate their tendency for automation-induced complacency (Cronbach's $\alpha = .714$). Participants indicated their agreement (on a 1–5 scale) with 20 statements created to indicate the likelihood

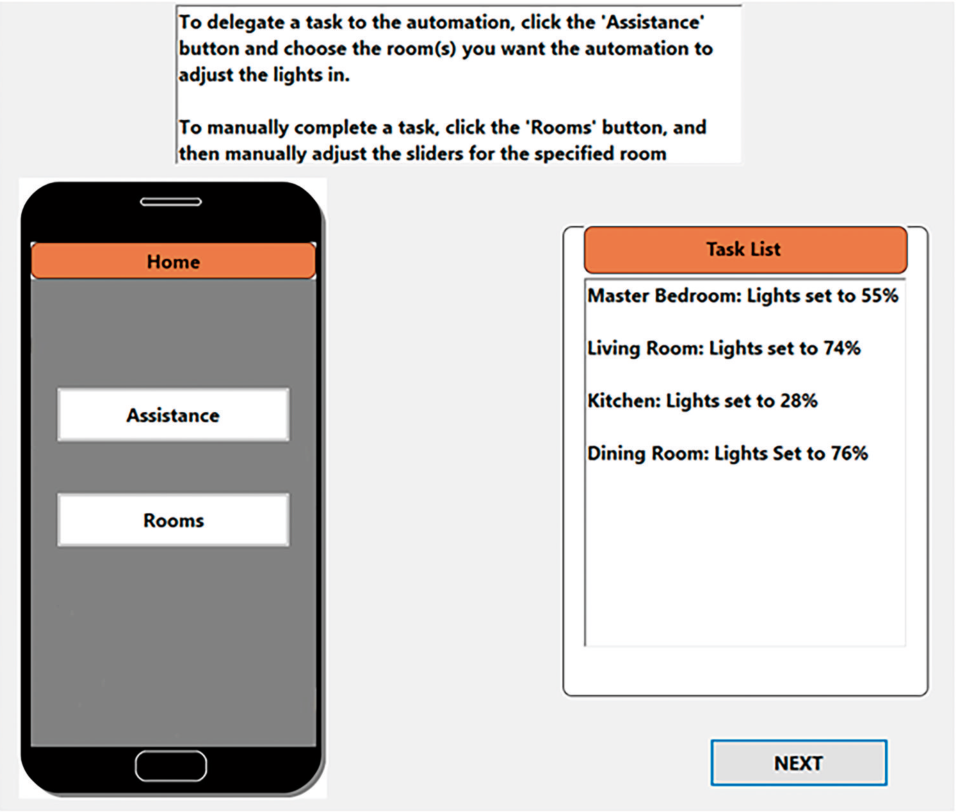

Figure 2. The main interface of the experimental task.

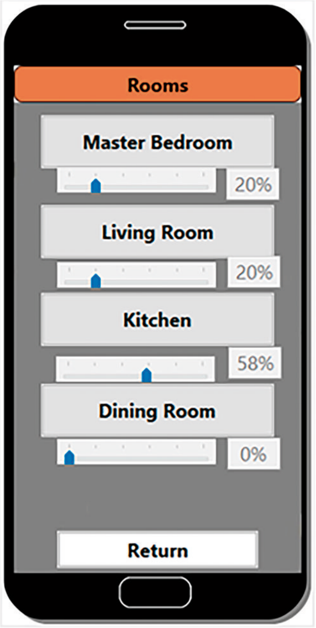

Figure 3. The 'Rooms' screen of the phone application.

of being complacent with common examples of automation (e.g. automated teller machines, ACC). CPRS scores were calculated by summing all agreement ratings, creating a range from 20 (low complacency potential) to 140 (high complacency potential). Trust was measured using one item, adapted from Lee and Moray (1992, 1994). Participants were required to self-report their trust in each subsystem. For example, they would report trust in the bedroom subsystem by responding to the item, *'To what extent did you trust (i.e. believe in the accuracy of) the automation aid to adjust the bedroom lighting in this scenario?'* Additionally, participants reported their self-confidence to complete the experimental task without the overall system by responding to, *'To what extent were you self-confident that you could successfully perform without the automation aid(s) in this scenario?'* Participants indicated the degree to which they agreed with the trust and self-confidence statements using a 0–100 visual analog scale, where higher scores indicate higher measures of the respective variable. Lastly, we measured participants' automation use and manual override behaviours. Automation use was measured as the proportion of trials when participants used a given subsystem. Manual override behaviours were measured as the proportion of trials a participant automated a task and subsequently manually completed the task on the 'Rooms' page.

Design and procedure

The experiment was a 2 (number of voiced assistants: one, four) × 2 (reliability: reliable, unreliable) mixed-

Figure 4. The feedback window. In this image, the participant automated the Dining Room's lightings and did not adjust the lights in the remaining rooms.

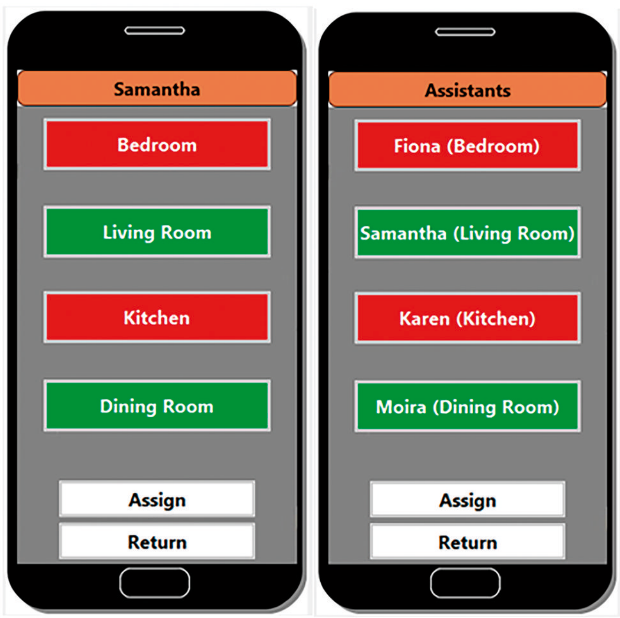

Figure 5. The 'Assistance' screen in the single-voiced (left) and four-voiced (right) conditions.

factors design, with reliability as the within-subjects variable. In the single voiced assistant condition, the 'Assistance' screen listed a single simulated voiced assistant that was responsible for adjusting the lights in all four rooms (Figure 5, left). In the four voiced assistants condition, the 'Assistance' screen listed four assistants, with one assistant assigned per room (Figure 5, right). We chose four voices from macOS's text-to-speech to create the voices: Samantha, Moira, Fiona, and Karen. We decided to use female voices because the most widely available smart assistants on the market tend to default to female agents (e.g. Amazon's Alexa, Google's Assistant, Apple's Siri). In the single-voiced condition, we randomly selected one of the voices to use as the single assistant that controlled all the rooms and interfaced with the participant. In the four assistants condition, each room was randomly assigned a single assistant to control the lighting.

In the reliable condition, all four room lighting subsystems were 100% reliable, meaning that no errors were present during these trials. In the unreliable condition, the living room lighting subsystem's reliability decreased to 60%, meaning that errors occurred for 40% of these trials. The remaining rooms (i.e. master bedroom, kitchen, and dining room) maintained 100% reliability. We chose 60% reliability because systems performing below 70% reliability are considered unreliable (Wickens and Dixon 2007; Kantowitz, Hanowski, and Kantowitz 1997). Participants completed the reliable condition first because errors early in trust formation are more damaging to trust than errors later in the interaction (Wickens, Helleberg, and Xu 2002). When the assistant failed to correctly perform an assigned task the assistant would say: *'I was not able to correctly adjust the lights in the living room.'* We chose to use non-apologetic neutral language to prevent any interpretations of trust repair. The way trust violations are addressed by the transgressor can have varying effects on how the violations impact trust (Kim et al. 2004), but the current study did not address trust violation remedies. When an assistant successfully completed a task, it would say: *'I was able to correctly adjust the lights in the* (room).*'* Note that an assistant only responded if the participant delegated a task to that assistant; that is, participants received solely text feedback for manually completed tasks.

The entire study duration varied between 50 and 60 minutes. First, each participant was randomly assigned to either the single- or four-voiced condition. After signing informed consent, participants were given verbal and written instructions for the task. Once participants confirmed they understood the instructions, they completed the CPRS questionnaire.

Table 1. Participant characteristics by agent-number condition (between-subjects factor).

	Single-agent				Four-agent			
	Male ($n = 18$)		Female ($n = 23$)		Male ($n = 21$)		Female ($n = 20$)	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Age	19.78	1.87	18.70	1.06	19.24	1.70	19.10	1.17
CPRS	100.33	10.00	100.35	9.83	99.43	10.68	96.10	8.64

Next, participants completed the reliable condition, which consisted of 12 trials. Then, participants completed the trust questionnaire to establish a baseline measure of trust. The participants then completed the unreliable condition, which consisted of 48 trials. Finally, participants completed the trust questionnaire a second time and were debriefed.

Results

Preliminary statistical tests

Prior to conducting our primary analyses, we tested to ensure that univariate distributions were normal and homoscedastic for the measures of trust, automation use, manual override behaviours, and self-confidence. Specifically, we performed Levene's Test for Equality of Error Variances, created Q-Q plots, and calculated skewness and Kurtosis metrics for each variable. The tests revealed that manual override behaviours exhibited floor effects at the first ($M = .09$, $SD = .20.$, skewness $= 3.306$) and second measurement points ($M = .06$, $SD = .19$, skewness $= 3.823$). Additionally, self-confidence ratings exhibited ceiling effects at the first ($M = 91.67$, $SD = 20.59$, skewness $= -3.215$) and second measurement points ($M = 93.35$, $SD = 18.77$, skewness $= -3.634$). That is, participants rarely manually adjusted sliders after delegating a task to an agent and were highly self-confident in their ability to complete the task without the automated system. Therefore, further analyses excluded manual override behaviours and confidence ratings. Participant characteristics of age (t (80) < 0.001, $p > 0.999$), gender (χ^2 (1, $N = 82$) $= .440$, $p = .507$), and CPRS scores (t (80) $= 1.175$, $p = .243$) did not significantly vary between agent conditions (see Table 1).

Tests for CST/SWT

The primary hypothesis relates to how the number of unique voices in a system influences participants' subjective trust and automation use. Therefore, we performed a 2 (number of voiced assistants: 1, 4) × 2 (reliability: reliable, unreliable) × 4 (room: living room, bedroom, kitchen, dining room) mixed doubly multivariate analysis of variance (MANOVA). We elected to first conduct a MANOVA instead of individual ANOVAs because of the possibility for multiple ANOVAs to inflate the familywise type I error rate (Tabachnick & Fidell, 2013, 286). The analysis revealed a significant three-way interaction (F (6,75) $= 3.199$, $p = .008$, $\eta_p^2 = .204$, see Table 2 for descriptive statistics). Next, we conducted follow-up univariate ANOVAs with trust and use to better explore how each variable is influenced by the same three independent variables.

Table 2. Trust and automation use by condition, subsystem (room), and reliability.

	Single-voiced ($n = 41$)				Four-voiced ($n = 41$)			
	Reliable		Unreliable[a]		Reliable		Unreliable[a]	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Trust								
LR	47.22	34.80	20.85	26.28	44.66	35.70	23.76	29.54
BR	45.24	35.93	67.80	35.15	46.05	36.30	41.88	40.92
K	46.49	35.26	64.51	36.60	45.17	36.36	42.80	41.39
DR	45.66	34.29	63.29	37.26	43.85	36.46	41.22	40.04
Use								
LR	0.35	0.32	0.28	0.28	0.35	0.33	0.22	0.27
BR	0.35	0.33	0.51	0.35	0.38	0.32	0.35	0.37
K	0.34	0.32	0.52	0.36	0.36	0.34	0.35	0.38
DR	0.34	0.32	0.50	0.36	0.35	0.32	0.34	0.37

Notes. LR: Living Room; BR: Bedroom; K: Kitchen; DR: Dining Room.
[a]In the unreliable condition, the living room subsystem was 60% reliable and the remaining subsystems were 100% reliable.

Trust

The first ANOVA revealed a main effect for room (F (3, 78) $= 19.587$, $p < .001$, $\eta_p^2 = .430$), and post hoc analyses revealed that trust ratings in the living room subsystem were significantly lesser than trust ratings in the remaining rooms (all $ps < .05$). However, trust ratings remained consistent for the singular variables of reliability condition (F (1,80) $= 0.003$, $p = .955$) and number of voiced assistants (F (1, 80) $= .918$, $p = .341$). The main effect of room was qualified by the presence of a significant three-way interaction (F (3, 78) $= 5.407$, $p = .002$, $\eta_p^2 = .172$). Post hoc analyses revealed that the source of the interaction is that participants in the single-voiced condition lost trust in the living room subsystem once it became unreliable, but they also gained trust in the other three rooms (all $ps < .05$, see Figure 6(A)). Participants in the four-voiced condition also lost trust in the living room subsystem ($p < .05$), but their trust in the other three rooms remained unchanged (all $ps > .05$, see Figure 6(B)). Considering that the unreliable living room subsystem influenced trust ratings for the other subsystems in the one-voiced condition but not in the four-voiced condition, it appears that the single-voiced

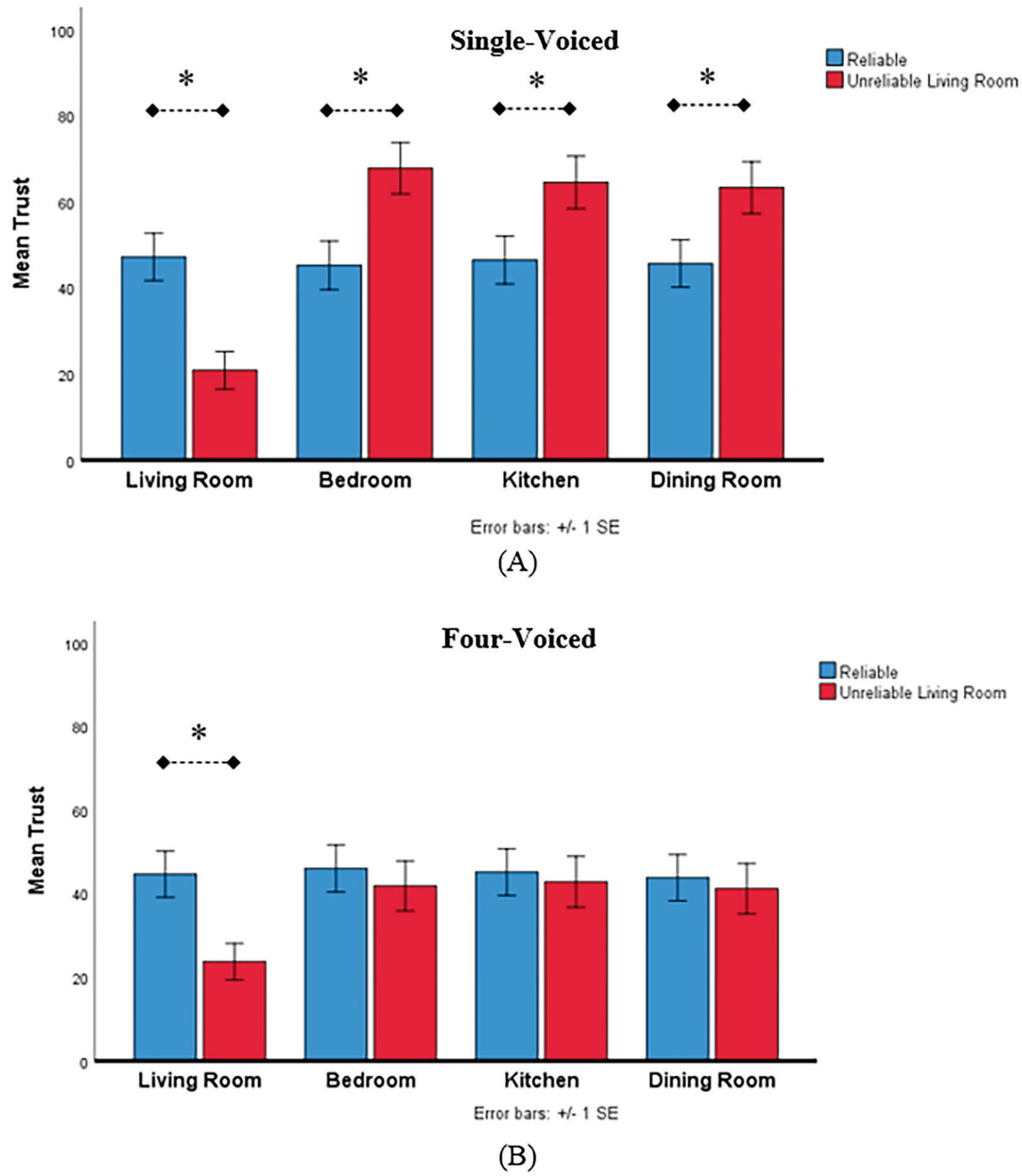

Figure 6. Mean trust comparisons between rooms, sorted by reliability. (A) Single-voiced system (B) Four-voiced system. Significant differences are indicated by an asterisk ($p < .05$).

manipulation may have promoted SWT, and the four-voiced manipulation may have promoted CST.

Automation use

Like trust, the second ANOVA and revealed a main effect for room (F (3,78) = 12.650, $p < .001$, $\eta_p^2 = .327$). Post hoc analyses displayed that participants automated the living room lighting significantly fewer times than the other rooms (all $ps < .05$). Additionally, there was a two-way interaction between room and reliability (F (1, 80) = 9.202, $p < .001$, $\eta_p^2 = .261$; see Figure 7). Post hoc analyses revealed that participants tended to increase their use of the reliable lighting subsystems (bedroom, kitchen, and dining room) and decrease their use of the living room subsystem after the living room became unreliable (all $ps < .05$). Contrary to our hypothesis, this interaction was unaffected by the number of voiced assistants in a system (F (3, 78) = 1.257, $p = .295$).

Discussion

In accordance with Nass, Steuer, and Tauber's (1994) findings, the results of this study suggest that multiple voices presented from a single system are perceived as individual actors or subsystems, evident in the trust differences between the single- and four-voiced

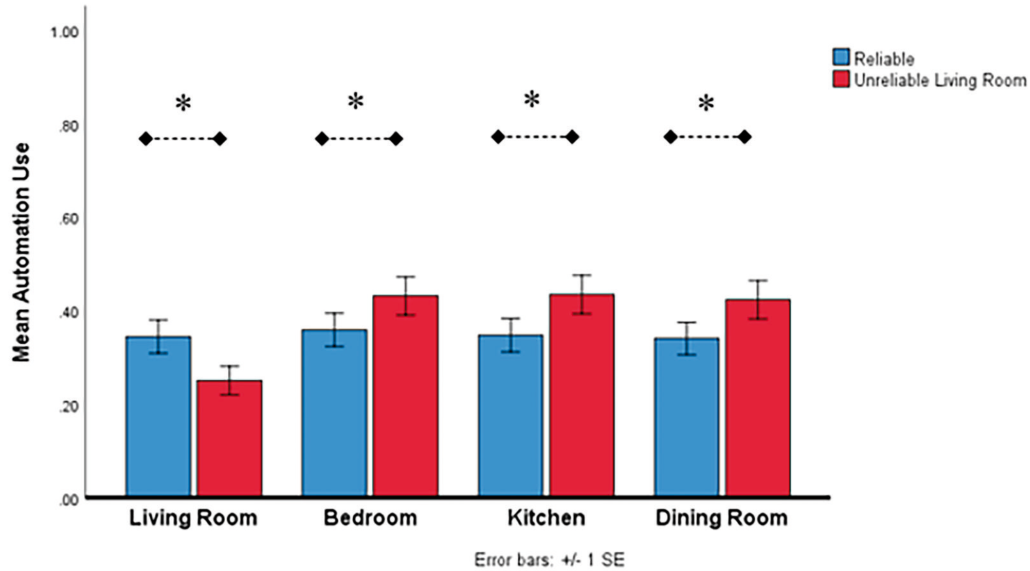

Figure 7. Mean automation use comparisons between rooms, sorted by reliability. Significant differences are indicated by an asterisk ($p < .05$).

conditions. Participants presented with a single-voiced system reported that changes in trust in the unreliable subsystem coincided with trust changes in reliable subsystems. Conversely, participants presented with a four-voiced system reported that their trust in reliable subsystems was consistent despite trust decreasing for the unreliable subsystem. Interestingly, participants in the single-agent condition reported increased trust in the reliable subsystems after the living room subsystem became unreliable. Although unexpected, prior work has found a similar effect where a reliable subsystem is considered more trustworthy once it is paired with an unreliable subsystem (Ross 2008). Additionally, we predicted a similar effect for automation use, but we found that automation use was similar between both conditions. Although unexpected, this finding does exclude differences in system exposure as a possible explanation for trust rating disparities between conditions. Instead, a likelier explanation for trust differences is that users of the single-voiced system adopted a SWT strategy while users of the more-easily distinguishable four-voiced system adopted a CST strategy. An alternative interpretation is that participants in the single-voiced condition better calibrated their trust to the true capabilities of the reliable subsystems; after all, trust values averaged around 50 during the reliable block, despite every subsystem being 100% reliable. The increase in trust for the reliable subsystems for the single-voiced condition could be explained simply as an attempt to calibrate their trust. While this interpretation has merit, this would ignore the trends showing that perceptions of a subsystem changed as its reliability remained consistent and the reliability of another subsystem decreased (a strong sign of SWT). A future study where reliability remained consistent across an extended period (like Lee and Moray (1992, 1994)) could explore this possibility of trust calibrating for each subsystem.

Another contribution of our findings is the extension of multi-component system literature findings to an automated system in the consumer domain. Compared to industrial- and military-oriented automation, consumer-oriented automation has less attentional demands and less severe consequences for failure (Pak et al. 2017). Our findings show that even in a simulated consumer-oriented automation task that (1) allowed users to choose whether to automate the task, (2) did not impose time pressures on the participants and (3) did not punish users for poor performance, participants could adopt either SWT or CST. However, the current study alone cannot determine any direct effect due to domain. Future studies can adopt an experimental design appropriate for comparing two similar tasks that only vary in domain (e.g. Guyton and Pak 2020).

While our manipulation of the number of voiced agents allowed us to promote different trust strategies, other factors may explain why previous studies overwhelmingly found evidence to support SWT as the dominant strategy. First, one factor that differentiates this study is the implementation of humanness design features for promoting component differentiability. Interface anthropomorphism has varying effects on user perceptions, including eliciting more social responses from users (Gong 2008) and variably influencing trust and system dependence (Pak et al. 2012). One possible rationalisation for our findings is that trust strategy adoption may be influenced by the level of humanness within a system. For example, a study that implemented a military target-identification task found that contagion effects differ between human duos and non-anthropomorphic machine duos (Ross 2008), which suggests that trust strategies vary according to the human likeness of each component within a team/system. This may explain why in previous studies that lack humanness design the unreliable subsystem 'pulls down' reliable subsystems (e.g. Keller and Rice 2009; Kluck et al. 2018), but in the current study, the unreliable subsystem increased participants' trust in the reliable subsystems. Research has yet to explore trust strategy differences across varying humanness levels in automated systems, but any potential effects may explain our results.

Another rationale for our findings is that previous studies that support SWT as the dominant strategy tended to grant users low decisional freedom (Lopez and Pak 2020). Decisional freedom reflects a user's flexibility in determining how to use an automated system (Hoff and Bashir 2015). SWT-supporting studies often present users with time constraints (e.g. Foroughi et al. 2019), a difficult primary task (e.g. Geels-Blair, Rice, and Schwark 2013), and/or a secondary task (e.g. Bean, Rice, and Keller 2011). Experimental designs that incorporate these elements can vastly increase workload and lead users to consider complete reliance on automation as the only viable option to maintain sufficient task performance (Biros, Daly, and Gunsch 2004). The current study allowed participants to either automate or manually complete the single experimental task at their own pace, granting participants a high level of decisional freedom. This claim is supported by the ceiling effect for self-confidence, the floor effect for manual override behaviours, and relatively low automation usage (less than 50% of trials). A likely explanation is that participants considered the task easy enough to test varying levels of manually completing the task without fear of performance decrements. Further, participants had enough confidence to automate a task and not override the system's behaviour, even for the unreliable subsystem. Considering these findings, one possibility is that a user's level of decisional freedom may also influence freedom in trust strategy adoption. That is, contexts that provide low decisional freedom may only afford SWT, whereas contexts that provide high decisional freedom allow SWT and CST. If this happens to be the case, then our findings may only extend to other contexts that afford high decisional freedom.

Another limitation to our study is our reliance on potentially novice system users. Some studies have recruited military cadets that may have some familiarity with the automated system (e.g. Foroughi et al. 2019), but most trust strategy studies recruit undergraduates with minimal system experience (Lopez and Pak 2020). The current study did not assess participants' level of familiarity with smart home systems, and total system interaction time was less than an hour. Most other studies provide participants with similar amounts of system exposure, with only three studies taking place over multiple days and sessions (Lee and Moray 1992, 1994; Lewandowsky, Mundy, and Tan 2000). These three are also some of the few studies to find evidence of users adopting CST. A yet untested variable is the effect of system exposure on trust strategy adoption. Users without system experience default to adopting SWT (e.g. Mehta & Rice, 2013), but this trend may change as the human-automation relationship develops. Additional research should monitor trust strategies as they change over time.

The current study used humanness design features in an attempt to influence the trust strategies adopted by users. Prior research in automated systems suggests that interface manipulations like the number of agents in a system should influence users to perceive components as independent in multi-agent systems. Our findings demonstrate that this single manipulation has effects on user perceptions (i.e. trust), but perhaps less effect on behaviours (i.e. automation use, manual override behaviours). Additional features that promote humanness to greater degrees (e.g. an embodied agent, systems that adapt to user behaviours, systems that incorporate human behaviours) may produce greater effects. The benefit of many of these humanness design methods, like the one used in this study, is that they can be implemented with minor changes to current systems. For instance, the manipulation used in the current study (the number of voiced assistants) can be added into any existing smart home systems without modifying the underlying workings of the system (i.e. the processes that would control the different aspects of the house). These results have direct implications for the potential design of future multi-component automated systems that are marketed towards consumers.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The author(s) reported there is no funding associated with the work featured in this article.

ORCID

Jeremy Lopez http://orcid.org/0000-0002-5451-5048
Richard Pak http://orcid.org/0000-0001-9145-6991

References

Bean, N. H., S. C. Rice, and M. D. Keller. 2011. "The Effect of Gestalt Psychology on the System-Wide Trust Strategy in Automation." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 55 (1): 1417–1421. doi: 10.1177/1071181311551295.

Beggiato, M., and J. F. Krems. 2013. "The Evolution of Mental Model, Trust and Acceptance of Adaptive Cruise Control in Relation to Initial Information." *Transportation Research Part F: Traffic Psychology and Behaviour* 18: 47–57. doi:10.1016/j.trf.2012.12.006.

Biros, D. P., M. Daly, and G. Gunsch. 2004. "The Influence of Task Load and Automation Trust on Deception Detection." *Group Decision and Negotiation* 13 (2): 173–189. doi:10.1023/B:GRUP.0000021840.85686.57.

Campbell, D. T. 2007. "Common Fate, Similarity, and Other Indices of the Status of Aggregates of Persons as Social Entities." *Behavioral Science* 3 (1): 14–25. doi:10.1002/bs.3830030103.

De Visser, E. J., R. Pak, and T. H. Shaw. 2018. "From 'Automation' to 'Autonomy': The Importance of Trust Repair in Human-Machine Interaction ." *Ergonomics* 61 (10): 1409–1427. doi:10.1080/00140139.2018.1457725.

Durso, F. T, and S. D. Gronlund. 1999. "Situation Awareness." In *Handbook of Applied Cognition*, edited by F.T. Durso, R.S. Nickerson, R.W. Schvaneveldt, S.T. Dumais, D.S. Lindsay, and M.T.H. Chi, 283–314. New York: John Wiley & Sons Ltd.

Endsley, M. R. 2017. "From Here to Autonomy: Lessons Learned from Human–Automation Research." *Human Factors* 59 (1): 5–27. doi:10.1177/0018720816681350.

Endsley, M. R., B. Bolte, and D. G. Jones. 2003. *Designing for Situation Awareness: An Approach to Human-Centered Design*. London, UK: Taylor & Francis.

Foroughi, C. K., C. Sibley, N. L. Brown, E. Rovira, R. Pak, and J. T. Coyne. 2019. "Detecting Automation Failures in a Simulated Supervisory Control Environment." *Ergonomics* 62 (9): 1150–1161. doi:10.1080/00140139.2019.1629639.

Fox, J. M. 1996. "Effects of Information Accuracy on User Trust and Compliance." In *CHI 1996 Conference on Human Factors in Computing Systems*, 35–36. New York: Association for Computing Machinery.

Geels-Blair, K., S. Rice, and J. Schwark. 2013. "Using System-Wide Trust Theory to Reveal the Contagion Effects of Automation False Alarms and Misses on Compliance and Reliance in a Simulated Aviation Task." *The International Journal of Aviation Psychology* 23 (3): 245–266. doi:10.1080/10508414.2013.799355.

Gong, L. 2008. "How Social is Social Responses to Computers? The Function of the Degree of Anthropomorphism in Computer Representations." *Computers in Human Behavior* 24 (4): 1494–1509. doi:10.1016/j.chb.2007.05.007.

Guyton, Z, and R. Pak. 2020. "The Impact of Automation Etiquette on User Performance and Trust in Non-Personified Technology." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 64 (1): 1805–1809. doi:10.1177/1071181320641435.

Hoff, K. A., and M. Bashir. 2015. "Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust." *Human Factors* 57 (3): 407–434.

Itoh, M. 2012. "Toward Overtrust-Free Advanced Driver Assistance Systems." *Cognition, Technology & Work* 14 (1): 51–60. doi:10.1007/s10111-011-0195-2.

Kantowitz, B. H., R. J. Hanowski, and S. C. Kantowitz. 1997. "Driver Acceptance of Unreliable Traffic Information in Familiar and Unfamiliar Settings." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 39 (2): 164–176. doi:10.1518/001872097778543831.

Kazi, T. A., N. A. Stanton, G. H. Walker, and M. S. Young. 2007. "Designer Driving: drivers' Conceptual Models and Level of Trust in Adaptive Cruise Control." *International Journal of Vehicle Design* 45 (3): 339–360. doi:10.1504/IJVD.2007.014909.

Keller, D., and S. Rice. 2009. "System-Wide Versus Component-Specific Trust Using Multiple Aids." *The Journal of General Psychology* 137 (1): 114–128. doi:10.1080/00221300903266713.

Kim, P. H., D. L. Ferrin, C. D. Cooper, and K. T. Dirks. 2004. "Removing the Shadow of Suspicion: The Effects of Apology Versus Denial for Repairing Competence-Versus Integrity-Based Trust Violations." *The Journal of Applied Psychology* 89 (1): 104–118.

Kluck, M., S. C. Kohn, J. C. Walliser, E. J. de Visser, and T. H. Shaw. 2018. "Stereotypical of Us to Stereotype Them: The Effect of System-Wide Trust on Heterogeneous Populations of Unmanned Autonomous Vehicles." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 62 (1): 1103–1107. doi:10.1177/1541931218621253.

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. doi: 10.1006/ijhc.1994.1007.

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 K. A. See. 2004. "Trust in Automation: Designing for Appropriate Reliance." *Human Factors* 46 (1): 50–80.

Lewandowsky, S., M. Mundy, and G. Tan. 2000. "The Dynamics of Trust: Comparing Humans to Automation." *Journal of Experimental Psychology. Applied* 6 (2): 104–123.

Lopez, J, and R. Pak. 2020. "Does One Bad Machine Spoil the Bunch?: A Review of Trust in Multiple-Component Systems." *Proceedings of the Human Factors and*

Ergonomics Society Annual Meeting 64 (1): 1546–1550. doi: 10.1177/1071181320641370.

Mehta, R., and S. Rice. 2016. "System Wide Trust: A Possible Contagion Effect." *Collegiate Aviation Review International* 34 (1):12–34. doi:10.22488/okstate.18.100465.

Moray, N. 1998. "Identifying Mental Models of Complex Human–Machine Systems." *International Journal of Industrial Ergonomics* 22 (4–5): 293–297. doi:10.1016/S0169-8141(97)00080-2.

Nass, C., J. Steuer, and E. R. Tauber. 1994. "Computers Are Social Actors." In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, 72–78. New York, NY: ACM.

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.

Pak, R., E. Rovira, A. C. McLaughlin, and N. Baldwin. 2017. "Does the Domain of Technology Impact User Trust? Investigating Trust in Automation Across Different Consumer-Oriented Domains in Young Adults, Military, and Older Adults." *Theoretical Issues in Ergonomics Science* 18 (3): 199–220. doi:10.1080/1463922X.2016.1175523.

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. doi:10.1518/001872097778543886.

Rice, S., and K. Geels. 2010. "Using System-Wide Trust Theory to Make Predictions About Dependence on Four Diagnostic Aids." *The Journal of General Psychology* 137 (4): 362–375.

Rice, S., S. R. Winter, J. E. Deaton, and I. Cremer. 2016. "What Are the Predictors of System-Wide Trust Loss in Transportation Automation?" *Journal of Aviation Technology and Engineering* 6 (1): 1. doi:10.7771/2159-6670.1120.

Ross, J. M. 2008. "Moderators of Trust and Reliance across Multiple Decision Aids." Doctoral diss., University of Central Florida.

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

Singh, I. L., R. Molloy, and R. Parasuraman. 1993. "Automation-Induced "Complacency": Development of the Complacency-Potential Rating Scale." *The International Journal of Aviation Psychology* 3 (2): 111–122. doi:10.1207/s15327108ijap0302_2.

Tabachnick, B. G, and L. S. Fidell. 2013. *Using Multivariate Statistics* (6th ed.). Boston, MA: Pearson Education Ltd.

Walliser, J. C., E. J. de Visser, and T. H. Shaw. 2016. "Application of a System-Wide Trust Strategy When Supervising Multiple Autonomous Agents." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 60 (1): 133–137. doi:10.1177/1541931213601031.

Wickens, C. D., A. S. Mavor, R. Parasuraman and J. P. McGee, eds. 1998. *The Future of Air Traffic Control: Human Operators and Automation.* Washington: National Academy Press.

Wickens, C. D., and S. R. Dixon. 2007. "The Benefits of Imperfect Diagnostic Automation: A Synthesis of the Literature." *Theoretical Issues in Ergonomics Science* 8 (3): 201–212. doi:10.1080/14639220500370105.

Wickens, Christopher D., John Helleberg, and Xidong Xu. 2002. "Pilot Maneuver Choice and Workload in Free Flight." *Human Factors* 44 (2): 171–188.