Not All Trust Is Created Equal: Dispositional and History-Based Trust in Human-Automation Interactions

Stephanie M. Merritt, University of Missouri–St. Louis, and Daniel R. Ilgen, Michigan State University

Objective: We provide an empirical demonstration of the importance of attending to human user individual differences in examinations of trust and automation use. Background: Past research has generally supported the notions that machine reliability predicts trust in automation, and trust in turn predicts automation use. However, links between user personality and perceptions of the machine with trust in automation have not been empirically established. Method: On our X-ray screening task, 255 students rated trust and made automation use decisions while visually searching for weapons in X-ray images of luggage. Results: We demonstrate that individual differences affect perceptions of machine characteristics when actual machine characteristics are constant, that perceptions account for 52% of trust variance above the effects of actual characteristics, and that perceptions mediate the effects of actual characteristics on trust. Importantly, we also demonstrate that when administered at different times, the same six trust items reflect two types of trust (dispositional trust and history-based trust) and that these two trust constructs are differentially related to other variables. Interactions were found among user characteristics, machine characteristics, and automation use. Conclusion: Our results suggest that increased specificity in the conceptualization and measurement of trust is required, future researchers should assess user perceptions of machine characteristics in addition to actual machine characteristics, and incorporation of user extraversion and propensity to trust machines can increase prediction of automation use decisions. Application: Potential applications include the design of flexible automation training programs tailored to individuals who differ in systematic ways.

INTRODUCTION

The past half-century has seen an unprecedented explosion of autonomous technology development and implementation. The widespread implementation of automated technologies in the workplace has led to a unique situation in which humans and machines must work jointly on the same task. This situation demands that users make decisions regarding how best to interact with automated systems. In such situations, the first decision that operators often must make is whether to use the automated system. This decision is a precursor to any influence the system may have on performance and therefore has important implications for the effectiveness of interaction strategies, yet it is not well understood. Because automation use decisions (and their resulting effects) can have life-and-death implications (Parasuraman & Riley, 1997), it is essential that researchers identify factors influencing them.

One major factor linked to automation use is trust. Although much research has examined machine performance as a predictor of trust, the role of user individual differences has received relatively little empirical attention. The purpose of the present study is twofold. First, we demonstrate that individual differences in personality (extraversion) and propensity to trust machines can increase prediction of automation use decisions.
measure of trust reflects two distinct types of trust when used at different times.

We begin by reviewing existing literature on the trust and automation use relationship. We next discuss two types of trust: dispositional trust and history-based trust. Although several types of trust have been identified (Kramer, 1999), these two types are particularly relevant to human-computer interaction and are thus our focus here. Dispositional trust reflects trust in other persons (or machines) upon initially encountering them, even if no interaction has yet taken place. In contrast, history-based trust is founded on interactions between the person and another person or machine. Thus, although dispositional trust represents a relatively stable construct, history-based trust is a dynamic construct that adjusts as a function of the person and machine’s cumulative interactions.

**Trust and Human-Computer Interaction**

Research on trust and human-computer interaction has generally supported two conclusions: (a) Individuals use machines that they trust more than those they do not trust, and (b) automation errors negatively affect trust. The relationship of trust and automation use was examined in a series of studies using variations of the PASTEURIZER task (Lee & Moray, 1992, 1994; Muir, 1989; Muir & Moray, 1996). In this task, participants oversee the operations of a virtual pasteurizer plant incorporating three pumps. In different variations of the task, participants can allocate control of one or more pumps to an automatic system. Muir and Moray (1996) found a positive correlation between trust and the extent to which users allocated control to the automated system ($r = .71$). Additional evidence supporting the relationship between trust and automation use has been provided by other researchers using the same and other paradigms (e.g., de Vries, Midden, & Bouwhuis, 2003; Dzindolet, Peterson, Pomranky, Pierce, & Beck, 2003; Johnson, 2004; Lee & Moray, 1994).

Several studies have also verified the common-sense proposition that trust is negatively affected by automation errors. Trust in the system was negatively affected by machine errors on a route-planning task (de Vries et al., 2003), a visual search task (Dzindolet et al., 2003), the PASTEURIZER task (Lee & Moray, 1992), and others (Johnson, 2004; Moray, Inagaki, & Itoh, 2000; Wiegmann, Rich, & Zhang, 2001).

It has been suggested that user individual differences may affect trust and, therefore, automation use decisions (e.g., Atoyan, Duquet, & Robert, 2006; Lee & See, 2004; Nickerson & Reilly, 2004; van Dongen & van Maanen, 2005). However, these propositions have generally appeared in theoretical work and have not been examined empirically. Some exceptions include work on self-confidence (e.g., de Vries et al., 2003; Lee & Moray, 1994), age (Wiegmann, McCarley, Kramer, & Wickens, 2006), and expectations about the utility of using the automation (Dzindolet, Pierce, Beck, & Dawe, 2002). In general, however, research on human-automation interactions has devoted more empirical attention to the characteristics of the automation than to those of the human.

Trust is generally thought to be affected by characteristics of the user, characteristics of the trustee (machine), and characteristics of the situation (Kramer, 1999). In the present study, we examine person and machine characteristics while holding the situation constant. We begin by discussing person characteristics, including dispositional trust and propensity to trust. We then discuss relevant trustee (machine) characteristics. Finally, we discuss how person and machine factors might interact to affect history-based trust.

**Characteristics of the User**

*Dispositional trust and propensity to trust machines.* Dispositional trust has been conceptualized as a diffuse expectation of others’ trustworthiness. It is based on early trust-related experiences and eventually becomes a stable personality characteristic (Hardin, 1996; Rotter, 1971, 1980). The implication is that individuals with a greater disposition toward trusting others will demonstrate greater levels of trust in another entity upon initial contact with that entity. In the present study, we distinguish between (a) the stable, trait-like tendency to trust or not trust others, which we call propensity to trust, and (b) the actual level of trust expressed in a specific other, which we call dispositional trust. We suggest that propensity to trust will be positively correlated with dispositional trust, but they are likely to be distinguishable constructs. Dispositional trust may be influenced by additional factors beyond propensity to trust, such as visible features of the other, as suggested by models of social impression formation (e.g., Fiske & Neuberg, 1990).

Authors have proposed that just as individuals have a general propensity to trust or distrust other
people, they may have a propensity to either trust or distrust machines (Atoyan et al., 2006; Muir & Moray, 1996; Nickerson & Reilly, 2004; Parasuraman, Molloy, & Singh, 1993). For example, automation-induced complacency potential may represent a general tendency to trust automation (i.e., propensity to trust machines; Singh, Molloy, & Parasuraman, 1993), which is likely to influence dispositional trust in a specific automated system.

In the present study, we assess trust in an automated system at two different points in time: after a brief introduction to the system (i.e., initial trust level) and after completion of a task in which the person and machine work jointly (i.e., posttask trust level). Although we measure trust with the same items in both cases, we suggest that the “trust” we measure initially is qualitatively different from the “trust” we measure posttask. We propose that the initial administration of our trust measure is most reflective of a dispositional trust construct because it is prior to extended experience with the automated system.

In contrast, when we administer the same measure following a period of interaction between user and machine, we believe the measure is most reflective of a history-based trust construct (which is conceptualized as a combination of the effects of dispositional trust and experience/interactions with the machine). Based on past research linking trust with automation use decisions (as described previously), we propose that initial levels of trust (which are measured after only a brief introduction to the machine and which we believe to represent a dispositional trust construct) will predict the extent to which participants use the system on a subsequent task.

Furthermore, dispositional trust levels (levels of trust in a specific other upon encountering that other) are thought to be strongly affected by propensity to trust. In contrast, history-based trust is thought to result largely from interactions with the machine (although propensity to trust may still have some effect on history-based trust). Therefore, if we are correct in suggesting that our initial trust measurement is more indicative of dispositional trust and our posttask trust measurement is more indicative of history-based trust, then stable propensity to trust machines should be more strongly associated with initial trust than posttask trust.

Hypothesis 1: Initial trust is positively related to automation use on a subsequent task.

Hypothesis 2: Propensity to trust machines is more strongly related to initial trust than to post-task trust.

Hypothesis 3: Initial trust mediates the relationship between propensity to trust machines and automation use.

Extraversion. It was suggested earlier that propensity to trust machines develops based on early trust-related experience. Thus, it is proposed to represent an experience-based construct that becomes relatively stable over time. Might there also be a truly dispositional, or personality-based, component of propensity to trust machines? Individuals high in extraversion tend to be sociable, assertive, and dominant, and they draw energy from being around others. In contrast, introverts tend to be quiet and reserved. They draw energy from being alone (McCrae & Costa, 1987). Research on interpersonal relationships has established that, on average, extraverts are more willing to trust other people than are introverts (Gaines et al., 1997; Omole & McLennan, 2000; Shikishima, Hiraishi, & Ando, 2006). However, we do not know whether extraverts are also more likely to trust nonpersons, such as machines. Because extraversion is a social trait, the relationship between extraversion and trust may not generalize to propensity to trust machines. On the other hand, claims often treat machines in social ways (Reeves & Nass, 1996). Therefore, the positive relationship between extraversion and propensity to trust may hold for machines.

Hypothesis 4: Extraversion is positively related to propensity to trust machines.

Characteristics of the Trustee (Machine)

Muir (1987, 1994) outlined trust-related factors for automated decision aids. Her theory relied heavily on two theories of human-human trust: Rempel, Holmes, and Zanna’s (1985) theory and Barber’s (1983) theory. Based on the trust factors described by Muir, we propose that four machine characteristics may affect trust in the automated system.

The first machine characteristic, competence, captures the machine’s ability to meet performance standards. In the case of an automated decision aid, a competent machine produces correct recommendations, whereas an incompetent machine does not (Lee & Moray, 1992). The second factor, responsibility, is the extent to which clear explanations of how the machine works (e.g., the algorithms it uses to reach decisions) are available to
the user (Muir, 1994). Providing information to the user is termed responsible in that it reduces the knowledge gap between the machine and the operator. Third, *predictability* refers to the extent to which the machine behaves consistently in a given situation. In the case of human-computer interaction, trust will be based on the consistency of the machine’s output given the same input. The final factor, *dependability*, focuses on the machine’s ability to consistently convert input to output over time. To the extent that the machine does not consistently perform (e.g., the machine “freezes up” or produces error messages), perceptions of dependability are compromised.

Empirical studies have linked these machine characteristics with perceptions of trust. Dzindolet et al. (2003) found that providing information regarding why the automation might err (i.e., increasing responsibility) was associated with increased trust. De Vries et al. (2003) found that machine competence was significantly related to ratings of trust. Muir and Moray (1996) found that competence was the best predictor, accounting for 81% of the variance in trust. Adding responsibility to the regression equation did not result in a significant improvement, although this may have resulted from a measurement artifact; the items assessing competence and responsibility were phrased almost identically. Predictability and dependability were analyzed in a separate regression model (therefore, they were not compared with competence and responsibility). Both predictability and dependability were significantly related to subjective trust. Similarly, Lee and Moray (1992), using separate regression analyses, found that the two factors accounted for 73.4% and 77.9% of the variance in trust, respectively. The evidence suggests both that the factors are related to trust and that they are highly intercorrelated; therefore, we analyze all factors simultaneously. Hypotheses related to machine characteristics are discussed in regard to history-based trust later.

**Interactive Effects of Trustor and Trustee**

In the following section, we propose several interactive effects among individual differences, machine characteristics, perceptions of those machine characteristics, and automation use. *Moderator effects* are those in which the level of the outcome variable depends on the levels of the two predictor variables. Hypotheses 8, 10, and 11 are moderator hypotheses. *Mediator effects* are those in which predictor variable A causes predictor variable B, which then causes outcome C. In this case, variable B is said to fully mediate the relationship between A and C. If A has a direct effect on C as well as the indirect effect it has via variable B, then variable B is said to partially mediate the relationship between A and C. Hypothesis 13 is a mediator hypothesis.

**History-based trust.** As previously discussed, history-based trust refers to the level of trust that exists based on one’s past interactions with the machine. This type of trust is continuously recalibrated to ongoing interactions, as each interaction provides additional information that can be used to make predictions about machine behavior (Kramer, 1999). According to this model, trust changes in the direction of experience (trustworthy or not trustworthy) and in an amount reflecting the degree of discrepancy between expected trustworthiness and encountered trustworthiness (Boyle & Bonacich, 1970). It is worthy of note that propensity to trust machines may have a lingering effect on history-based trust; thus, history-based trust can be viewed as a composite of one’s stable tendency to be trusting and one’s actual experiences with the other. Most research on human-automation trust has focused on the role of machine characteristics and therefore considers history-based trust.

The idea that various types of trust exist at different stages of human-automation interaction has been considered. According to Lee and See (2004), initial trust is thought to relate to affective processes, whereas if the user decides to rely on the system, subsequent trust is thought to be more heavily influenced by analytic processes related to perceptions of machine performance. In addition, Atoyian et al. (2006) proposed that different scales of human-automation trust are more appropriate at different levels of experience. However, researchers have not yet empirically demonstrated a distinction between early and later forms of trust. We seek to do so here.

We measure trust after only 1 minute of exposure to an automated system (initial trust), and we reassess trust after completion of a 20-minute task (posttask trust). If we are correct that initial and posttask measurements of trust reflect two different types of trust, then these two measurements should not correlate so highly that they would be considered to represent the same construct. However, because we believe that history-based trust
is influenced by both experiences and dispositions, we expect the two measures of trust to be moderately positively correlated. We therefore expect the following:

Hypothesis 5: Initial trust in an automated system is positively associated with posttask trust.
Hypothesis 6: Initial trust and posttask trust show evidence of discriminant validity.

As discussed previously, several studies have established links between the machine’s characteristics and trust; therefore, we expect that machine characteristics will be associated with trust measured posttask. Because interaction with the machine is an essential component of history-based trust, but not dispositional trust, machine characteristics should be more strongly associated with history-based trust than with dispositional trust. If we are correct that initial trust largely represents dispositional trust, whereas posttask trust largely reflects history-based trust, then machine characteristics should be more strongly associated with posttask trust than with initial trust (although because participants have had a minimal exposure to the machine prior to rating initial trust, machine characteristics may also be positively related to initial trust). An alternative hypothesis is that once users become aware of the machine’s characteristics via their minimal exposure, the machine characteristics have an immediate and constant effect on trust. If this alternative hypothesis is correct, machine characteristics should be equally related to initial and posttask trust.

Furthermore, we suggest that trust measured at a given point in time may lie on a continuum between dispositional and history-based trust. We propose that as the user interacts with the machine over time, trust is likely to move along the continuum from heavily dispositional to heavily history-based. However, this shift is contingent on actual interaction with the machine. If operators disuse the system, they are unable to observe or have experiences with its characteristics. In this case, posttask ratings of trust may remain heavily influenced by dispositions but are not likely to relate to the machine’s characteristics. If our proposal is correct, we should find an interaction between propensity to trust machines and automation use, such that posttask trust is more strongly affected by the machine’s characteristics when individuals use the machine more often.

Hypothesis 7: Machine characteristics are more strongly related to posttask trust than to initial trust.

Hypothesis 8: Automation use moderates the relationship of machine characteristics and posttask trust, such that machine characteristics have a greater effect on posttask trust when use is high.

Person × Machine interaction effects. Based on their individual differences, different people do not always perceive the same situation in the same way (Terborg, 1981). Thus, individual perceptions of the machine characteristics may vary even when the objective machine characteristics are held constant. The notion that user perceptions of automation reliability may differ from actual reliability is incorporated into the theoretical models proposed by Dzindolet and colleagues (Dzindolet et al., 2002; Dzindolet, Pierce, Beck, Dawe, & Anderson, 2001), Lee and See (2004), and Madhavan and Wiegmann (2007). Relatedly, some researchers have reported seemingly significant differences in perceptions of machine reliability when actual reliability was fixed (e.g., Dzindolet et al., 2003; Johnson, 2004; Madhavan, Wiegmann, & Lacson, 2003, 2006; Wiegmann et al., 2001). However, correlates of that variance have yet to be examined empirically.

Variance in posttask trust may be related to the interaction between propensity to trust machines and machine characteristics. In cases dealing with social situations, Robinson (1996) found that psychological contract breaches were perceived less severely for individuals with higher initial trust in their employers. Furthermore, when psychological contract breach was perceived, employees with lower initial trust in their employers experienced a greater decline in trust than did individuals with higher initial trust. Generalizing from Robinson’s findings would suggest that individuals with lower propensity to trust machines will react more negatively to a poorly performing machine than will individuals with higher propensity to trust machines.

However, the opposite hypothesis is suggested by Boyle and Bonacich (1970), who proposed that the decline in trust resulting from poor interaction occurs in proportion to the discrepancy between the prior expectation for trustworthiness and the actual trustworthiness. In that case, individuals with higher propensity to trust machines would react more negatively to a poorly performing machine than would individuals with lower propensity to trust machines. Because conflicting
perspectives are evident in the literature, we propose a nondirectional moderation hypothesis.

Hypothesis 9: Significant variance in perceptions of machine characteristics exists even when actual machine characteristics are constant.

Hypothesis 10: Machine characteristics moderate the relationship of propensity to trust machines and posttask trust.

A more complex view of the process would be that perceptions of machine characteristics may arise from the combination of objective machine characteristics and propensity to trust machines. Expectations for a machine’s performance may guide us to perceive ambiguous events as consistent with our expectations (e.g., Dzindolet et al., 2003). Therefore, individuals with a greater expectancy for machine performance (i.e., high propensity to trust machines) should be more likely to notice and process positive aspects of machine performance and vice versa. We expect that the highest levels of perceived machine characteristics will be observed in cases where both propensity to trust machines and machine characteristics are positive; conversely, the lowest levels will be found when both propensity to trust machines and machine characteristics are low.

Hypothesis 11: Machine characteristics will moderate the relationship of propensity to trust machines and perceptions of machine characteristics, such that more extreme levels of posttask trust will result when both variables are either low or high.

Finally, we propose that because individuals base their behavior on their perceptions of the world around them, their perceptions of the machine characteristics may be closer (i.e., more proximal) predictors of history-based trust than the objective machine characteristics. We therefore propose the following:

Hypothesis 12: Perceptions of machine characteristics account for additional variance in posttask trust beyond the effects of the actual machine characteristics.

Hypothesis 13: Perceptions of machine characteristics mediate the relationship between the interaction of propensity to trust with actual machine characteristics and posttask trust.

METHOD

Overview: The X-Ray Screening Task

We designed a task that would be somewhat familiar to our participants and that would be seen as important — the task of inspecting luggage for dangerous items. The X-Ray Screening Task is a medium-fidelity, computer-based simulation in which X-ray images of suitcases are presented on a monitor. Some images contain weapons; most do not. Participants inspect each image as quickly and accurately as possible and indicate whether they would “search” the bag (they believe it contains a weapon) or “clear” the bag (they believe it contains no weapon). Participants are also told that an automated machine is available to assist them. In the present study, this (fictitious) machine was called the Automatic Weapons Detector (AWD). A screenshot from the X-Ray Screening Task is displayed in Figure 1.

Our X-Ray Screening Task can be described as what Lee and See (2004) call a “microworld,” a simplified version of a real-life setting that retains...
essential elements but enables increased experimental control (Lee & See, 2004, p. 65; see also Brehmer & Dorner, 1993). The primary essential elements that we wished to incorporate into our microworld task are outlined as follows.

**Competing demands.** Users of automated technology often face competing demands on their time and cognitive resources. The primary responsibility of airport screeners is to prevent dangerous items from entering the aircraft; however, overly thorough searches would result in long lines, missed flights, and customer dissatisfaction. They must therefore balance the competing demands of speed and accuracy. To simulate this trade-off, the X-Ray Screening Task rewards both speed and accuracy with points accumulated.

**Motivation and familiarity.** The timely nature of aviation security is likely to induce intrinsic motivation. In addition, because many Americans are already familiar with the nature of airport security screening, they are likely to learn the task with relatively little training. To further minimize the need for training, we limited the “weapons” category to guns and knives. Including other prohibited but less easily identifiable items (e.g., explosives) would have increased the amount of training required.

**Situation strength.** Past research has sometimes created strong situations that controlled automation use, leaving little chance for individual differences to show effects (Lee & Moray, 1992). It was important that the X-Ray Screening Task present a sufficiently weak situation. We controlled situation strength in three ways. First, we set the AWD to provide correct recommendations at rates similar to that of the average participant. If the AWD’s accuracy were set too close to 100% or below chance, the optimal strategy (i.e., use/nonuse) would be obvious, and use behaviors might be driven simply by obvious strategy choices rather than individual differences (but see Beck, Dzindolet, & Pierce, 2007; Dzindolet et al., 2002).

Second, we imposed a point-scoring system on performance. A scoring system that rewarded the identification of weapons too strongly relative to the costs of searching a “safe” bag would create a situation in which the optimal strategy would be obvious — search every bag. We thus created a scoring system in which, based on the percentage of weapons present and the points awarded for correct and incorrect decisions, either searching or clearing every image resulted in a loss of points.

Third, consistent with conditions found outside the laboratory, using the AWD was associated with a cost in terms of speed. Each time the AWD was activated, we imposed a 5-second time delay that we told participants was “processing time.” Because overall performance was dependent on both speed and accuracy, use of the AWD thus came with a cost. This cost contributes to the weakening of the situation such that participants are discouraged from adopting a simple strategy to use the AWD every time.

**Participants and Design**

A sample of 258 undergraduate students from a large midwestern university received course credit for participation. Participants’ mean age was 19.25 years, and the sample was 84.9% White, 4.7% Black, 4.7% Asian American, and 3.1% other ethnicities (2.4% declined to report ethnicity). To increase motivation, the top scorer in each condition received a $50 prize. Scores were based on both speed and accuracy of performance. Three participants’ simulation task data were lost as a result of technical difficulties; thus, our total sample for the variables assessed during and after the simulation task (i.e., use, perceived machine characteristics, final trust) was 255.

**Procedure**

Experimental sessions were held in an on-campus computer lab. After providing informed consent, participants completed an online questionnaire assessing control variables, propensity to trust machines, and extraversion. Next, participants completed a short training session that included instructions on what to look for (guns and knives), how to “search” or “clear” a bag, and the scoring system, which was pilot tested prior to the study (see Table 1). After training, participants completed a 1-minute practice trial during which they did not have access to the AWD (i.e., they screened several bags without assistance).

Participants next received training on how to use the AWD. This training included descriptions of the AWD’s competence, predictability, and dependability (to be discussed further in the Experimental Manipulations section) and a 1-minute demonstration during which participants witnessed the AWD’s tendency to err (competence) and/or break down (dependability). During this instruction and 1-minute demonstration, participants were
able to form a first impression of the machine's characteristics. They then rated their initial levels of trust in the AWD.

After the two training sessions, participants were given 20 minutes to screen as many bags as possible while minimizing errors. The base rate for weapon presence was 30%. Although this percentage is high compared with the real-world rate of guns and knives in airport luggage, screeners are also required to identify more common prohibited items, such as scissors or liquids. Furthermore, this hit rate is similar to the precedent set in past laboratory research (Dzindolet et al., 2001). The difficulty of the X-ray slides varied, with unassisted performance ranging from a low of 5% correct to a high of 100% correct. In general, easier slides contained fewer items, and if there was a weapon present, it was easily recognizable. Difficult slides contained many overlapping items, and if a weapon was present, it was presented at an angle that made it more difficult to recognize.

On average, 76% of unassisted participants were correct on each slide. Overall performance (assisted and unassisted) averaged 3,330.10 points with a standard deviation of 984.57 (three scores for outliers greater than 3 standard deviations from the mean were removed). Immediately following each “search” or “clear” decision, the participant viewed feedback indicating whether the decision was correct or incorrect. Following the task period, participants completed the final trust measure.

### Experimental Manipulations

Machine characteristics were manipulated to create two (high and low) machine function conditions, and within each condition, each participant experienced the same levels of machine competence, responsibility, predictability, and dependability. The manipulations are described as follows and in Table 2.

#### Competence
In the high machine function condition, participants were told that the machine would make correct recommendations 85% of the time. In the low machine function condition, participants were told that the AWD was correct only 65% of the time. These percentages were pilot tested for effectiveness and were reflected in the AWD’s actual functioning throughout the trials. In terms of signal detection, the AWD’s $d’$ was 2.01 in the high-function condition and .49 in the low-function condition. In both cases, $c$ was smaller than $-0.01$, indicating almost no bias toward search or clear recommendations.

#### Responsibility
Responsibility was not manipulated verbally but was reflected in the amount of information provided during the 5 seconds required for the AWD to operate. In the high condition, three messages were displayed: “acquiring,” “scanning,” and “processing.” In the low condition, a single message, “scanning,” was displayed for the entire delay.

#### Predictability
In the high machine function condition, participants were told that the AWD was 100% predictable – that is, if given the same bag to screen repeatedly, the AWD would make the same recommendation each time. Participants in the low machine function condition were told

### Table 1: X-Ray Screening Task Scoring System

<table>
<thead>
<tr>
<th>Action</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correctly identify a gun or knife</td>
<td>+100</td>
</tr>
<tr>
<td>Miss a gun or knife</td>
<td>-100</td>
</tr>
<tr>
<td>Correctly clear a bag</td>
<td>+25</td>
</tr>
<tr>
<td>Open a bag and find nothing</td>
<td>-25</td>
</tr>
<tr>
<td>Each minute elapsed</td>
<td>-10</td>
</tr>
</tbody>
</table>

### Table 2: Machine Characteristics Manipulations

<table>
<thead>
<tr>
<th>Factor</th>
<th>High</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>Competence</td>
<td>Correct on 85% of trials</td>
<td>Correct on 65% of trials</td>
</tr>
<tr>
<td>Responsibility</td>
<td>Provides three status messages: “acquiring,” “scanning,” and “processing”</td>
<td>Provides one status message: “scanning”</td>
</tr>
<tr>
<td>Predictability</td>
<td>Produces same recommendation when it sees the same bag multiple times</td>
<td>May produce different recommendations when it sees the same bag multiple times</td>
</tr>
<tr>
<td>Dependability</td>
<td>Does not break down</td>
<td>Breaks down on 25% of trials</td>
</tr>
</tbody>
</table>
that the AWD might produce different recommendations when the same bag is scanned repeatedly.

**Dependability.** In the high machine function condition, participants were told that the AWD does not break down. Participants in the low machine function condition were informed that the AWD may fail, and during the task, the AWD produced error messages 25% of the time, on average, when the AWD was used.

**Measures**

**Propensity to trust machines.** Propensity to trust machines was assessed using 12 items from the Automation-Induced Complacency Potential Rating Scale (Singh et al., 1993), which presents scenarios involving different machines. Example items include the following: “If I need to have a tumor in my body removed, I would choose to undergo computer-aided surgery using laser technology because it is more reliable and safer than manual surgery” and “Automated systems used in modern aircraft, such as the automatic landing system, have made air journeys safer.” The response scale is a 5-point Likert-type scale ranging from strongly disagree to strongly agree. In previous research, the scale had an internal consistency reliability estimate of $\alpha = .90$, and this scale has been previously discussed as a measure representing the propensity to trust machines (Lee & See, 2004).

**Extraversion.** Extraversion was measured using the International Personality Item Pool (IPIP) measure (Goldberg, 2001). The scale consists of 10 items on a 5-point Likert-type scale ranging from very inaccurate to very accurate. Example items include the following: “I make friends easily” and “I am the life of the party.” Gow, White, Pattie, and Deary (2005) found that this scale had internal consistency reliabilities ranging from .84 to .90 in three samples.

**Initial and posttask trust.** Both initial and posttask trust were measured using the same six-item scale, which was created for this study. One item assesses overall trustworthiness of the AWD, and the other five items each relate to the trust-related factors in Muir’s (1987) theory. Example items include the following: “Overall, I think the AWD is trustworthy” and “The AWD is a competent performer.” Respondents answered on a 5-point Likert-type scale ranging from strongly disagree to strongly agree.

**Automation use.** Automation use was assessed with the proportion of total bags that were screened using the AWD. A bag was considered screened using the AWD if the participant activated the AWD prior to making the search/clear decision. Thus, automation use represents the decision to receive advice from the AWD.

In regard to automation use, we note that in our study, participants on average demonstrated high rates of AWD disuse. This was inconsistent with the results from our pilot sample but consistent with results from previous work (e.g., Beck et al., 2007; Dzindolet et al., 2002). To increase variance on automation use, we instituted a mandatory use condition for half the participants, in which the AWD automatically provided a recommendation for each slide. Thus, 131 participants were able to choose whether to activate the AWD; 124 participants had no choice.

**Perceived machine characteristics.** Fifteen self-report items were administered following the task for the dual purpose of (a) checking the effectiveness of the machine characteristics manipulation and (b) measuring perceptions of machine characteristics. These items assessed participants’ perceptions of predictability (3 items; e.g., “In general, the AWD does the same thing each time I use it”), dependability (4 items; e.g., “I am certain that the AWD will not break down”), responsibility (3 items; e.g., “The way the AWD works is a mystery to me”), and competence (5 items; e.g., “The AWD usually gave me the correct recommendation”). These items were adapted from the trust scale by Rempel et al. (1985) or written by the researchers. Because machine characteristics were constant within condition, we expected that the average score for perceived machine characteristics would be higher in the high machine characteristics group than in the low machine characteristics group; however, consistent with Hypothesis 9, we expected variance on these scales—even within condition.

**RESULTS**

**Results Overview**

Before turning to the results, we note that, as mentioned earlier, a mandatory use condition was created for half of the experimental sessions. Independent samples $t$ tests revealed no significant differences between participants in the optional and mandatory use conditions on any predictor variables. Hypotheses 1, 3, and 8, which concerned
automation use, were tested using the AWD-optional participants only ($N = 131$). All other hypotheses were tested using the full sample ($N = 255$).

**Descriptive Statistics**

Scale means, standard deviations, alpha reliabilities, and intercorrelations are shown in Table 3. All scales had acceptable reliability with the exception of propensity to trust machines. This scale previously had shown high levels of reliability (Singh et al., 1993), but those levels failed to replicate. The cause is unclear, as both our study and the Singh et al. (1993) study used student samples. To improve the scale’s reliability, we dropped five items with low item-total correlations. This increased the scale’s reliability to $\alpha = .65$. A unidimensional confirmatory factor analysis using the remaining seven items fit the data closely ($\chi^2 = 13.50$, $df = 13$, $ns$, root mean square error of approximation [RMSEA] = .01, confirmatory fit index [CFI] = 1.00, nonnormed fit index [NNFI] = 1.00); therefore, we proceeded with hypothesis testing using the reduced seven-item scale.

**Control Variables**

In the analyses to follow, we controlled for age, sex, and ethnicity to rule out the possibility that our results were caused by these factors. We also assessed participants’ self-reported knowledge about computers and comfort with using computers (assessed with one item each) to assess whether the three demographic factors were related to computer comfort and knowledge. The three demographic factors were used simply as controls; however, the correlations between these demographic variables and those of substantive interest are included in Table 3 for interested readers.

**Manipulation Check**

Independent samples $t$ tests on the perceptions of machine characteristics showed that the means of the high and low machine characteristics groups were significantly different in all cases ($p < .01$), with the differences in the expected direction (see Table 4).

Hypothesis 1 proposed that initial trust would be positively related to automation use on the task. Using only the participants for whom automation use was optional, we performed a linear regression in which the control variables (age, sex, and ethnicity) were entered into Block 1 and initial trust was entered into Block 2. The relationship between initial trust and automation use was significant ($\beta = .27, p < .01$), supporting Hypothesis 1.

Hypothesis 2 proposed that because we suggest that the initial trust measurement represents dispositional trust and the posttask trust measurement represents history-based trust, propensity to trust machines should be associated more strongly with initial trust than with posttask trust. A structural equation model was fitted, in which propensity to trust machines predicted both initial and posttask trust. This model fit well ($\chi^2 = 194.45$, $df = 138$, RMSEA = .04, CFI = .98, NNFI = .98), and the path from propensity to trust machines to initial trust was significant ($\beta = .21, p < .05$), whereas the path from propensity to trust machines to posttask trust was nonsignificant ($\beta = -.07, ns$).

This model was contrasted with a second model, in which the paths from propensity to trust machines to each trust outcome were constrained to be equal. The constrained model also showed acceptable fit ($\chi^2 = 203.40$, $df = 139$, RMSEA = .04, CFI = .98, NNFI = .98), but the decrease in fit was statistically significant ($\Delta \chi^2 = 8.95$, $\Delta df = 1$, $p < .05$), demonstrating that propensity to trust machines was more strongly associated with initial trust than with posttask trust. These findings support Hypothesis 2 and our proposition that the initial measurement of trust was more reflective of dispositional trust than was posttask trust (which we propose represents history-based trust).

Hypothesis 3 proposed that initial trust mediates the relationship between propensity to trust machines and automation use. This hypothesis was tested using the Sobel (1982) test with a bootstrapping procedure, which does not assume a normal distribution and is therefore more appropriate for small samples (Preacher & Hayes, 2004). The indirect effect of propensity to trust machines on use was significant ($p < .05$), and full mediation was found, supporting Hypothesis 3. Hypothesis 4 suggested that extraversion is positively related to propensity to trust machines. A linear regression revealed a significant relationship ($\beta = .20, p < .01$), supporting Hypothesis 4.

Hypotheses 5 and 6 proposed that initial and posttask trust are positively related (Hypothesis 5), yet show evidence of discriminant validity (Hypothesis 6). Supporting Hypothesis 5, the correlation between initial and posttask trust was positive and significant ($r = .42, p < .05$). To test Hypothesis 6, we contrasted the fits of two confirmatory
### TABLE 3: Scale Means, Standard Deviations, Reliabilities, and Intercorrelations

<table>
<thead>
<tr>
<th>Scale</th>
<th>M</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>2.26</td>
<td>1.28</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Sex (male/female)</td>
<td>1.70</td>
<td>.46</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Ethnicity (majority/minority)</td>
<td>1.84</td>
<td>.37</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Computer comfort</td>
<td>4.40</td>
<td>.92</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Computer knowledge</td>
<td>3.36</td>
<td>.78</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Extraversion</td>
<td>3.75</td>
<td>.52</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Propensity to trust machines</td>
<td>3.50</td>
<td>.37</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Initial trust</td>
<td>2.90</td>
<td>.62</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Posttask trust</td>
<td>2.50</td>
<td>.76</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Automation use</td>
<td>.16</td>
<td>.19</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Correlations corrected for attenuation as a result of unreliability displayed in the upper half. Reliabilities for multi-item measures are displayed in the diagonal. Correlations between automation use and all variables are based on an n of 131, and all other correlations are based on the combined sample, with ns ranging from 253 to 258. Male and majority coded 1; female and minority coded 2.

* *p < .05.
factor analysis (CFA) models. In the first, the correlation between initial and posttask trust was freely estimated. This model fit was acceptable ($\chi^2 = 147.06$, $df = 53$, RMSEA = .09, CFI = .97), and the correlation between initial and posttask trust was significant ($r = .47$, $p < .05$), further supporting Hypothesis 5. In the second model, the correlation between the two trust factors was fixed to 1.0. This model fit somewhat worse ($\chi^2 = 191.38$, $df = 54$, RMSEA = .10, NNFI = .94, CFI = .95), and the chi-square difference test was significant ($p < .01$), indicating that the correlation between initial and posttask trust was significantly less than 1.0. These results support Hypothesis 6 and are consistent with our suggestion that the initial and posttask measurements of trust reflect two different types of trust.

Hypothesis 7 suggested that because we propose that the initial trust measurement represents dispositional trust and the posttask measurement represents history-based trust, the machine’s characteristics should be more strongly related to posttask trust than to initial trust. We performed a contrast similar to the one used to test Hypothesis 2. The baseline model ($\chi^2 = 91.91$, $df = 53$, RMSEA = .06, CFI = .99, NNFI = .98), in which the paths from machine characteristics to initial trust ($\beta = .28$, $p < .05$) and posttask trust ($\beta = .52$, $p < .05$) were both positive and significant, was compared with a constrained model. The overall fit decreased ($\chi^2 = 109.34$, $df = 54$, RMSEA = .07, CFI = .98, NNFI = .98), and this decrease was statistically significant ($\Delta\chi^2 = 17.43$, $\Delta df = 1$, $p < .05$). Machine characteristics were significantly stronger predictors of posttask trust than initial trust, supporting Hypothesis 7 and the suggestion that the posttask measurement of trust represents history-based trust to a larger extent than does the initial trust measurement.

Taken together, the findings of Hypotheses 2, 5, 6, and 7 suggest that initial levels of trust (which we propose represent dispositional trust) are more strongly influenced by stable propensity to trust, whereas later levels of trust (which we propose represent history-based trust) are more strongly influenced by characteristics of the machine itself. In other words, assessments of trust taken at different times seem to represent two different qualitative forms of trust, one of which is more heavily influenced by characteristics of the human and the other of which is more heavily influenced by characteristics of the machine.

Hypothesis 8 suggested that machine characteristics and automation use interact to predict posttask trust. A hierarchical linear regression was used. The direct effects of machine characteristics were significant ($p < .01$), and the direct effect of automation use was marginally significant ($p = .08$). The addition of the interaction increased the variance accounted for by a marginally significant amount ($p = .07$), and the effect of the interaction term was marginally significant ($p = .07$). The form of the interaction indicated that, as expected, when automation use was low, there was no difference in trust between participants in the high and low machine characteristics conditions. However, when use was high, participants in the high machine condition had much higher trust than did participants in the low machine condition, as displayed in Figure 2. The model accounted for 28% of the variance in posttask trust.

Hypothesis 9 proposed that significant variance exists in perceptions of machine characteristics within the machine characteristics condition (i.e., when actual machine characteristics were held constant). In a confirmatory factor-analytic approach, the variances of the latent perceptions constructs were examined. Separate models were constructed for each of the two actual machine characteristics conditions, so that actual machine characteristics were held constant.

We examined the significance of the phi parameters, which represent the variance of the latent
perceptions factors. In both cases, the variance of perceived competence and predictability was not significantly different from 0. In other words, within condition, all participants perceived the competence and predictability of the AWD similarly. The variance of dependability was significant ($p < .05$) in both cases, indicating that within condition, there were significant differences in the way the machine’s dependability was perceived. The variance of responsibility was significant only in the high condition ($p < .05$), indicating that there were significant within-condition differences in the way participants perceived the responsibility of the machine in the high condition, but in the low condition, participants’ perceptions of responsibility were relatively uniform. Overall, Hypothesis 9 was partially supported.

Hypothesis 10 suggested that the interaction of machine characteristics and the propensity to trust machines predicts posttask trust. A hierarchical linear regression supported the hypothesis. The direct effect of machine characteristics was significant ($\beta = .50, p < .01$), and the interaction term was also significant ($\beta = .56, \Delta R^2 = .01, p < .05$). The form of the interaction indicated that when the machine characteristics were high, participants with high propensity to trust had higher posttask trust than did participants with low propensity to trust. However, when machine characteristics were low (i.e., the machine functioned less well), the relationship reversed, such that participants with lower propensity to trust had higher posttask trust than did participants with higher propensity to trust machines (see Figure 3).

Hypothesis 11 proposed that the interaction of propensity to trust machines and machine characteristics predicts perceptions of machine characteristics. Four separate hierarchical regressions were run with the perceptions of the four machine characteristics as the dependent variables. For predictability and responsibility, the direct effects of machine characteristics were significant ($\beta = .46$ and .29, respectively, $p < .01$), but the interactions were not.

For dependability and competence, the direct effects of machine characteristics were again significant ($\beta = .46$ and .56, respectively, $p < .01$), and the interactions were significant. For perceived dependability, the interaction term was significant ($\beta = .57, p < .05; \Delta R^2 = .01, p < .05$). The form of the interaction indicated that participants in the high condition perceived higher dependability than participants in the low condition, but the difference was greater for participants with high propensity to trust machines. For perceived competence, the interaction was marginally significant ($\beta = .48, p = .06; \Delta R^2 = .01, p = .06$), and the form of the interaction was the same as that found for perceived dependability.

Hypothesis 12 suggested that perceptions of machine characteristics account for additional variance in posttask trust over and above the effects of actual machine characteristics. In a hierarchical linear regression, the effects of the actual machine characteristics were significant ($\beta = .51, p < .01$). In support of our hypothesis, the addition of the perceptions resulted in a large increase in variance accounted for ($\Delta R^2 = .52, p < .01$). Overall, the model accounted for 78% of the variance in posttask trust.

Finally, Hypothesis 13 proposed that perceptions of machine characteristics mediate the relationship...
between the interaction of propensity to trust with actual machine characteristics and posttask trust. The significances of the indirect effects were tested using the Sobel technique with bootstrapping, similar to the test for Hypothesis 3. In all cases, the indirect effect estimates were significant ($p < .01$), indicating that perceptions of competence, responsibility, predictability, and dependability significantly mediate the effects of the interaction of actual machine characteristics and propensity to trust on posttask trust. For perceived competence, the mediation effect was full. For perceptions of responsibility, predictability, and dependability, the effect was partial.

**DISCUSSION**

The results of the present study provide empirical evidence on several important points regarding trust and human-automation interaction. First, different trust constructs exist, and the different constructs show different relationships with correlates. Therefore, without theoretical attention to these distinctions, empirical relationships between predictors and outcomes may be muddled.

Second, we found that user propensity to trust machines and machine characteristics interact to affect subsequent trust ratings. Consistent with theory on history-based trust, we found an interaction between propensity to trust and automation characteristics such that people who had a higher propensity to trust but who were paired with an ill-functioning machine suffered the largest negative effects on perceptions of machine characteristics and on posttask trust. This finding is consistent with the schema approach to automation errors proposed by Dzindolet and colleagues (Dzindolet et al., 2002, 2003). Individuals with higher propensity to trust machines and higher initial trust are likely to expect the machine to perform correctly; in other words, they are likely to have a schema for automation that involves reliable performance. Machine outcomes inconsistent with this schema (i.e., errors) are more likely to be noticed and remembered by individuals who hold the schema more strongly. In turn, subsequent perceptions of machine characteristics and trust ratings are likely to be affected more severely for individuals who experience greater schema violation.

We also found a significant interaction between machine characteristics and the degree of use. As the rate of use increases, history-based trust levels tend to become more heavily influenced by machine characteristics and less influenced by propensity to trust machines.

Our third contribution was to empirically demonstrate the importance of user perceptions of automation. Inclusion of perceptions increased the variance accounted for in history-based trust by 52% over and above the variance accounted for by the actual machine characteristics. The variance in perceptions was predicted by the interaction of the actual machine characteristics and user propensity to trust, suggesting that although individuals do perceive the general level of machine functioning, their perceptions are also affected by their individual differences. Finally, these perceptions play an important mediating role between user and machine characteristics and history-based trust.

**Implications**

The results of this study have implications for both research and practice. In terms of research, the results indicate that a dynamic perspective on trust is required. Trust evolves over time, from dispositional trust to history-based trust. Therefore, researchers should measure trust at multiple points in time to gain a clearer picture of the differential impact that variables may have on early and later trust.

Furthermore, our results clearly imply that researchers who wish to determine the impact of machine characteristics on trust (or on behavioral/performance outcomes related to trust) should at the very least control for propensity to trust machines in their analyses. More effectively, researchers should test for interactions between individual differences, including propensity to trust machines and automation design features, as suggested by our results showing that propensity to trust interacts with machine characteristics in affecting trust.

Finally, our results clearly imply that researchers should increase the attention devoted to user perceptions of machine characteristics. Simply assessing the actual characteristics of the machines is not enough. Our study demonstrates that the same machine will be perceived differently by different users, and those perceptions mediate the relationship of machine characteristics with trust. Thus, to more accurately predict, explain, and control user behavior, researchers must assess user perceptions and determine the factors that contribute to individual differences in perceptions of the same machine.
Regarding practical implications, the major lesson learned from the present study is that different people may need different interventions to successfully calibrate their levels of trust. Because optimal trust levels require that the user’s perceptions of machine characteristics reflect the actual machine characteristics (Madhavan & Wiegmann, 2007), it seems likely that to encourage optimal strategy adoption, one should encourage the formation of history-based trust, rather than a reliance on dispositional trust. Thus, training design should consider both (a) the characteristics of the machine and (b) the user’s individual differences, and training should be customized to meet the individual user’s needs.

For example, individuals high in extraversion and propensity to trust machines are likely to have high levels of initial trust in a machine. If the machine is high functioning, these levels are likely to be reinforced – the danger inherent in this situation is that trust may become overly inflated. On the other hand, if the machine underwhelms these individuals’ expectations, trust is likely to drop severely, perhaps below the optimal level. Thus, customized information should be presented to combat such miscalibrations.

Training may be customized in several ways. One method might be to conduct a pretraining assessment, and trainees could be grouped for sessions according to their level of extraversion or (more effectively) propensity to trust machines. A second method might be to use computer-based adaptive training, in which technology is employed to provide each trainee with a customized program specifically tailored to his or her needs. Several potentially important training outcomes exist above and beyond knowledge and skills necessary to operate the system (cf. Ford, Kraiger, & Merritt, in press; Kraiger, Ford, & Salas, 1993). We suggest that practitioners may benefit from making proper calibration of trust (an attitudinal learning outcome) an explicit and valued objective of training programs.

Most implications suggest customization to different personalities; however, one implication of the present study holds across all individuals. During training relevant to the implementation of a new system, all trainees should be required to regularly use the system. Our results show that the posttask trust levels for individuals with high rates of use were more heavily influenced by characteristics of the machine, whereas for individuals with low rates of use, posttask trust continued to be heavily influenced by propensity to trust machines. We therefore recommend that, until trust is properly calibrated to the machine’s characteristics, all operators be required to consistently use and interact with the system.

Limitations and Suggestions for Future Research

We present several limitations and suggestions for future research. First, we used a sample of college students, which is common in past research. However, students may differ from employees in motivation levels. We attempted to increase motivation by offering a $50 prize to the top scorer in each condition. Participant comments indicated that this prize was an effective incentive. However, future researchers might examine the extent to which our results generalize to other populations. A second limitation was that the propensity to trust machines scale did not achieve an acceptable reliability level. Although this scale had previously demonstrated high reliability (Singh et al., 1993), we achieved a maximum alpha of .65. The cause of this replication failure is unclear. Thus, future researchers might wish to validate the findings of our study using other measures.

Third, our procedure manipulated all of the machine characteristics in the same direction simultaneously. This was consistent with past research indicating that many or all of these machine characteristics might be highly correlated (Lee & Moray, 1992; Muir & Moray, 1996). However, we were unable to examine whether the characteristics are differentially associated with trust. Furthermore, when the characteristics are examined individually, the particular manipulations of the characteristics used are likely to have increased importance. For example, our manipulation of responsibility was rather weak, and studies replicating our manipulation may find smaller effects for responsibility than might studies using a stronger manipulation. Thus, future researchers might compare the effects of various manipulations on outcomes such as trust, use, and reliance. Future researchers might also wish to determine whether specific manipulations may interact with user individual differences in affecting such outcomes.

In addition, future researchers might examine whether differences in weightings exist for rating trust in a machine versus trust in another human.
Our finding that extraversion is positively associated with propensity to trust machines implies that propensity to trust other humans and propensity to trust machines may be positively associated. However, research indicates that some different reactions to human versus machine aids may exist (Dzindolet et al., 2002; Lewandowsky, Mundy, & Tan, 2000; Madhavan & Wiegmann, 2007) or that trust in another human is influenced by a greater number of factors than trust in a machine (e.g., Bonaccio & Dalal, 2006; Madhavan & Wiegmann, 2007). Identification of similarities and differences in reactions to human and machine aids is an interesting avenue for future research.

Another fruitful avenue for future research might lie in measure development. Future researchers might develop trust measures that have clear discriminant validity for history-based, dispositional, and other relevant forms of trust. Researchers might also wish to empirically establish which of the many proposed forms of trust between humans (Kramer, 1999) can be reasonably applied to discussions of trust in machines. Conversely, researchers might wish to generate a trust measure that can be appropriately used to assess and compare levels of trust at any point along the continuum ranging from dispositional to history-based trust. Such scales should show appropriate properties of measurement invariance for assessments made initially and at later points in time (cf. Vandenberg & Lance, 2000).

Furthermore, future researchers should examine a larger array of user individual differences that might affect automation use or disuse decisions. Past research has established task self-efficacy (i.e., self-confidence) as a potentially important factor (de Vries et al., 2003; Lee & Moray, 1994). Self-confidence might parallel trust in that it may incorporate dispositional (e.g., self-esteem) and history-based (e.g., task self-efficacy) components. Future researchers might investigate this possibility.

Also, individuals might also disuse automation because of a need for control or a desire to hone or maintain one’s task-related skills (i.e., mastery goal orientation; van Dongen & van Maanen, 2005), and use may be related to openness and agreeableness personality characteristics (Nickerson & Reilly, 2004). Identification of relevant user individual differences is an important step toward building a comprehensive model of human-computer interactions.

Finally, an interesting avenue for future research would be to determine whether interactions might exist between individual differences in propensity to trust machines and aspects of system interface design. Might interface be customized to the individual user’s level of propensity to trust machines, so that perceptions of machine characteristics could be more accurately calibrated to reflect actual characteristics? Relatedly, researchers might vary the type of feedback provided to participants and assess its effects. For example, might users high in propensity to trust machines be more likely to discount feedback indicating the machine has erred when that feedback comes at a significant delay than when it is immediate? Furthermore, some tasks may provide asymmetric feedback—for example, airport security screeners receive immediate feedback for “search” decisions (i.e., either they find a prohibited item or they do not); however, feedback for “clear” decisions is usually delayed and indirect (i.e., did any adverse events occur on a flight?). Immediate feedback may be a motivating factor in itself, leading individuals to prefer engaging in actions that result in immediate feedback rather than those that do not. Researchers might wish to identify individual differences that may lead users to prefer actions that result in immediate feedback (e.g., tolerance of ambiguity or goal orientation).

**ACKNOWLEDGMENTS**

We acknowledge the Transportation Security Administration for providing X-ray images. We also thank Daniel Nadeau, Marcy Schafer, and Chip Shank for their valuable assistance.

**REFERENCES**


route planning. International Journal of Human-Computer Studies, 58, 719–735.


Stephanie M. Merritt is an assistant professor of industrial-organizational psychology at the University of Missouri–St. Louis. She received her Ph.D. in psychology from Michigan State University in 2007.

Daniel R. Ilgen is the John A. Hannah Distinguished Professor of Industrial-Organizational Psychology and Management at Michigan State University. He earned his Ph.D. in psychology from the University of Illinois in 1969 and is former editor of Organizational Behavior and Human Decision Processes and past president of the Society of Industrial and Organizational Psychology.

Date received: September 7, 2007
Date accepted: February 16, 2008