An Empirical Model of Cultural Factors on Trust in Automation

Shih-Yi Chien, Michael Lewis
School of Information Sciences
University of Pittsburgh
Pittsburgh, PA, 15260 U.S.A.
shc56@pitt.edu, ml@sis.pitt.edu

Zhaleh Semnani-Azad
Department of Psychology
University of Waterloo
Waterloo, Ontario, N2L3G1 Canada
zsemnani@uwaterloo.ca

Katia Sycara
Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15213 U.S.A.
katia@cs.cmu.edu

Abstract— Trust is conceived to be an important factor mediating an individual's reliance on automation. Studies have shown individual and cultural differences as well as tasking context significantly affect an individual's development of trust behaviors. This paper reports preliminary progress in developing a psychometrically grounded subjective measure of trust in automation. A total of 110 items from 8 existing instruments were considered for inclusion in this instrument using Amazon Mechanical Turk to supply samples. Exploratory factor analysis was performed to determine the dimensionality of the data, with 42 items selected for continued refinement. Our proposed model comprises 3 main constructs (performance expectancy, process transparency, and purpose influence) along with 3 types of moderators (cultural-technological contexts, individual, and cultural differences).

INTRODUCTION

The use of autonomous systems to reduce risk and enhance the efficiency of humans has become ubiquitous in modern society. The automation itself consists of various complex systems and sub-systems that in spite of the best design efforts may not work perfectly under all situations. Therefore, the beneficial effects of using automation (e.g., delivering more accurate information, lowering operator workload, or allowing the operator to make faster decisions) may not be realized due to maladaptive use of the automation. In human-automation systems, it has been observed that the human may fail to use the automation properly (i.e. abuse, disuse, or misuse the automation) or devote insufficient resources to monitor the automation (i.e. over-reliance). In particular, Parasuraman and Riley (1997) have defined misuse as overreliance (sometimes also called complacency) on automation (e.g. using the automation when it should not be used, failing to monitor it effectively), disuse as underutilization or under-reliance (e.g. ignoring or turning off alarms), and abuse as inappropriate application of automation by managers or designers. Misuse has contributed to many accidents through operator failure to monitor the automation in aviation (Funk, Lyall, & Wilson, 1999), marine navigation (Parasuraman & Riley, 1997) and other areas (Casey & Casey, 1993). Disuse can decrease automation benefits and lead to accidents, if, for instance, safety systems and alarms are turned off or ignored (Parasuraman & Riley, 1997).

Trust has frequently been cited (Lee & Moray, 1992) as a contributor to human decisions about monitoring and use of automation. In addition, trust is also relevant to automation bias (misuse or disuse), which affects both naïve and expert users. Meyerson, Weick, & Kramer (1996), for example, found that even skilled subject matter experts had misplaced trust in the accuracy of diagnostic expert systems. Additionally the Aviation Safety Reporting System contains many reports from pilots of incidents involving failure to monitor automated systems such as autopilots (Mosier, Skitka, Heers, & Burdick, 1997). Studies have shown that individual and cultural differences can significantly affect the development of an individual's trust behaviors.

For example, national culture as measured by Hofstede's cultural dimensions (Hofstede, 1991) were shown to exert a meaningful influence on attitude and behavior in human interaction with automation, even in a highly specialized and regulated profession (Merritt, 2000). On the other hand, a human's history of interaction with automation affects future behavior indirectly through dynamic changes in trust (i.e. individual differences in trust in automation). Studies such as (Jian, Bisantz, & Drury, 2000) have shown that cultural factors affect trust formation and degree of trust in authority, which have been cited in the trust in automation literature as possible causes of misuse and automation bias. Most of the existing works on cultural effects on trust in automation have been done in the context of interpersonal trust (Fulmer & Gelfand, 2010; Gunia, Brett, Nandkeolyar, & Kamdar, 2011). Additionally, most of the limited work studying culture and trust in automation failed to support by empirical validation.

The purpose of this research is to develop a psychometrically grounded measure that can be used to gain a greater understanding of general principles and factors pertaining to trust in automation. In future work this validated measure will be used along with a specially designed "trust sensitive task" to investigate how trust mediates reliance on automation across cultures. In this initial phase, we seek to construct a reliable psychometric instrument that captures the nature and antecedents of trust in automation across cultures. In the following sections, the development of a trust assessment instrument is reported.

METHOD

Initial Instrument Development

We began developing a standardized measure for trust in automation by pooling items from 8 existing measures, Culture-Technology Fit (CTF) by (Lee, Choi, Kim, & Hong, 2007), Empirically Derived (ED) by (Jian et al., 2000), Human-Computer Trust (HCT) by (Madsen & Gregor, 2000), International Comparison of Technology Adoption (ICTA) by (Im, Hong, & Kang, 2011), Online Trust Belief (OTB) by (Hwang & Lee, 2012), SHAPE Automation Trust Index (SATI) by (Goillau & Kelly, 2003), Technological Adoptiveness Scale (TAS) by (Halpert, Horvath, Preston, Somerville, & Semnani-azad, 2008), and Trust in Specific Technology (TIST) by (Mcknight & Carter, 2011), that have benefited from systematic development and validation.

In present form the items of the scale are based on questions regarding human participants'

(a) General attitudes toward automation, i.e. items involving predisposition to trust, which were adapted from ED, TAS, and TIST (e.g., I am confident in an automation/ I believe that most automations are effective at what they are designed to do).

(b) Attitudes invoked after human participants had been cued to think about particular instances of automation (such as an automated navigation/GPS aid), items were adapted from HCT, SATI, ICTA, OTB, and TIST (e.g., The automation uses appropriate methods to reach decisions).

(c) Attitudes across cultural-technological contexts (such as uncertainty avoidance and subjective norms), items were adapted from CTF, ICTA, and OTB (e.g., The senior management of this business/school has been helpful in the use of automation).

Classification of Items Purpose

Rather than being a unitary concept, the antecedents of trust in automation and subsequent automation use include a number of facets, with at least two common elements: general and specific uses of automation (Parasuraman & Riley, 1997). Our initial step was to categorize the characteristics of the population of items into the general or specific cluster. 45 participants were recruited from the University of Pittsburgh community to identifying the referents of the selected items. The participants were informed by the following instruction:

"Please respond to the following statements about your trust in automation. By automation, we mean any technology or service that you have used before, including apps, devices, functions, or systems. Based on your experience, use the following scale to rate the extent to which you agree (5) or disagree (1) with the statements below. Note, no wrong responses to any of the statements, the most critical is to record your own true opinion on each item. *If you think the provided instruction is **not sufficient** to answer a question, please rate it as insufficient information*."

A 5-point Likert scale (from strongly disagree to strongly agree) was adopted in the initial test, however, participants were allowed to rate an item as "**insufficient**" if they felt the item was too closely associated with a specific automated system for general use. Among 110 selected items, 70 items were identified as addressing automation in general, whereas 40 items involved judgments about particular instances of automation.

Scale Refinement – Exploratory Factor Analysis

To refine the population of items, an initial pilot test was conducted using Amazon Mechanical Turk (MTurk). A total of 65 paid participants (32% male and 68% female, average age= 38.1) were recruited on MTurk to complete the trust instrument, which consisted of the 110 items for measuring attitudes toward automation in general and specific use of automation. Smart phones and a navigation system (GPS) were chosen as the instances of general and specific automation respectively (Table 1).

TABLE 1. Types and Descriptions of Uses in Automation

Type	Description
General Auto	By "Automation" we mean any technology or service that takes actions automatically and that you have used, including apps, devices, functions, or systems.
Specific Auto	By "Automation" we focus mainly on GPS Navigation System including all types of navigation devices that you have used, such as an automotive navigation system (e.g., Garmin) or Smartphone navigation apps (e.g., Google map).

Exploratory factor analysis (EFA) was used to determine the dimensionality of the data while a principal components factor analysis with varimax rotation was performed to examine the number of factors produced. 59 questions met the validity criteria, 40 general (details not included due to the space limits) and 19 specific items (Table 4). A five-factor model was returned which explained 52.4% of the variance in answers to items addressing general automation (Table 2); whereas 70.2% of the variance in answers to items addressing specific instances of automation were captured by a five-factor model, however, as one of the factors failed to pass the internal consistency and reliability test (resulting Cronbach's alpha is lower than 0.7), that factor was eliminated (Table 3).

TABLE 2. Reliability Statistics in General Automation.

General Auto	Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	Num of Items
Factor 1	.922	.922	13
Factor 2	.867	.871	6
Factor 3	.892	.890	9
Factor 4	.863	.870	7
Factor 5	.718	.732	5

TABLE 3. Reliability Statistics in Specific Automation.

Specific Auto	Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	Num of Items
Factor 1	.943	.944	10
Factor 2	.836	.835	3
Factor 3	.818	.821	3
Factor 4	.626	.628	2
Factor 5	.760	.797	3

These data gathered in the test were used to refine our scale by rewording or identifying problematic items. After eliminating redundant items, 26 general and 16 specific items remained for use in constructing the proposed trust model and validation in a second round of data collection.

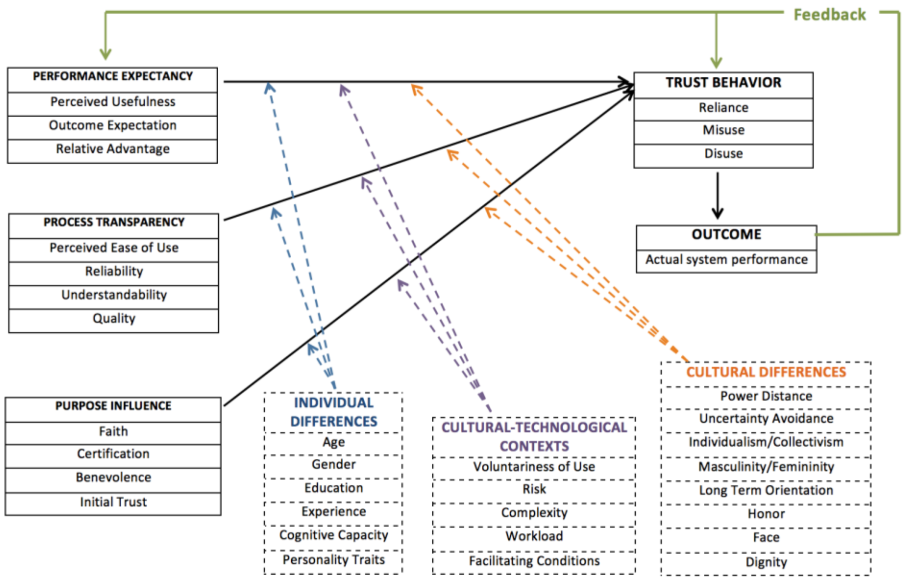

Figure 1. Model of factors of trust development in automation. Performance expectancy, process transparency, and purpose influence are the constructs (solid lines); individual differences, cultural-technological contexts, and cultural differences are the moderators (dotted arrows).

Model Development

To capture the direct and indirect effects of trust on specific situations, the proposed model (figure 1) includes three main constructs similar to those identified by Lee and See (2004) along with three types of moderators. These are expected to interact in complex ways to produce trust mediated behavior. This version of the instrument is comprised of 26 general and 16 specific items, categorized into three constructs (performance expectancy, process transparency, and purpose influence) and one moderator (cultural-technological contexts). The systematic constructs closely resemble those of Lee and See's model (Lee & See, 2004), in which they compared fourteen relevant measures and found most to involve only 2-3 dimensions. Their proposed model based on 3 dimensions designated performance, process, and purpose. This three dimensional structure fits nicely with Mayer et al.'s (1995) ability-integrity-benevolence definition that has been widely adopted in social psychological studies of trust and suggests that candidate items for an instrument measuring trust in automation should contain at least these dimensions.

Additionally, two more types of moderators were proposed to capture individual and cultural differences to better predict trust behaviors. We hypothesize these moderators will affect the main constructs and therefore indirectly cause changes in trust behaviors. An individual's trust could influence her reliance on automation as well as the system performance. Despite the system outcome, the individual might reevaluate her strategy in interacting with the automation. The following sections describe the model's constructs and moderators in greater detail.

TABLE 4. EFA (exploratory factor analysis) results: specific use of automation (21 items/5 factors). The values represented the factor loadings for each item. The model of specific items with a threshold value 0.4, in order to eliminate the noise. Note. (r): recode values.

Rotated Component Matrix[a]

	Component				
	1	2	3	4	5
I can rely on automation to ensure my performance.	.830				
Automation improves my performance.	.750				
It is easy to follow what automation does.	.725				
Automation makes use of all the knowledge and information available to produce its solution to the problem.	.708				
The advice automation produces is as good as that which a highly competent person could produce.	.696				
Automation correctly uses the information I provided.	.680				
I am confident in automation.	.646				
Automation always provides the advice I require to make my decision.	.627				
Using automation increases my productivity.	.606				
Using automation enables me to accomplish tasks more quickly.	.542				
Automation has sound knowledge about the type of problem for which it is intended.		**.726**			
Automation is friendly to use.		**.702**			
Automation uses appropriate methods to reach decisions.		**.618**			
Automation may result in unpredictable situations. (r)			.838		
Automation does not fail me.			.783		
I believe automation could be faulted. (r)			.761		
I understand how automation works.				.827	
I am wary of automation. (r)				.611	
I am suspicious of automation's intent. (r)					**.842**
Automation is deceptive. (r)					**.758**
Automation behaves in an underhanded manner. (r)					**.750**

Performance Expectancy

Performance expectancy is defined as an individual's belief that applying automation will help her to enhance job performance. The degree of trust will be affected by the past and consequences of system performance. 8 general and 4 specific items involved this cluster along with 3 dimensions: perceived usefulness, outcome expectancy, and relative advantage.

Perceived Usefulness

Perceived usefulness refers to an individual believes that using a particular system would assist her to achieve the goal. For example, a person may feel automation is useful on her tasks or using automation makes her tasks easier.

Outcome Expectation

Outcome expectancy relates to the belief that by receiving assistance from a system, an individual believes his job performance would be enhanced.

Relative Advantage

Relative advantage compares the differences in a user's preferences between interacting with another individual and relying on a particular instance of automation. For instance, a person may accept system predictions rather than the recommendations from a group of consultants.

Process Transparency

The transparency of automation may affect an individual's degree of perceived difficulty in using it (i.e., how it functions). 12 items (5 general and 7 specific) were adopted, distributed among 4 constructs: perceived ease of use, reliability, understandability, and quality.

Perceived Ease of Use

Perceived ease of use reports an individual's perceived cost or effort in learning and using an instance of automation to perform a job.

Reliability

The reliability of assistance would directly influence the decision to use automation. The failure rate, for example, may influence an individual's willingness to rely on a particular automation.

Understandability

Understandability refers to difficulties in comprehending how automation performs tasks and in predicting the outcome and consequence.

Quality

The quality of provided information might affect a person's trust in automation. For example, if the automation fails to provide sufficient information, an individual may ignore recommendations and switch off the automation.

Purpose Influence

Purpose influence relates to a person's knowledge of what the automation is supposed to do. 7 general and 5 specific items from the conducted study were involved in this cluster and they cover 4 dimensions: faith, certification, benevolence, and initial trust.

Faith

Faith refers to an individual's belief in future behavior of an instance of automation. For instance, people may rely on the recommendation from automation rather than herself when unsure about a decision.

Certification

The presence of a certification or product guarantee may lead to less worry about its potential flaws. For example, a third party seal would be critical for online banking systems.

Benevolence

Benevolence refers to an individual's beliefs that automation is designed with good intentions and will not diminish their performance.

Initial Trust

Initial trust refers to a person's instinctive tendency of trust when using an innovation. An individual may give it benefit of doubt when she first uses it.

Moderators

It is generally believed that trust is dynamic and interacts with a variety of other influences to determine behavior. A self-confident operator, for example, may operate a system manually because he believes he can "do a better job." An operator from a culture with high uncertainty avoidance may choose manual operation to avoid a slight possibility of automation error. Although the observed disuse of automation is the same in both cases, its cause and potentially effective interventions are very different. To assure that the measurement instruments are reliable across various contexts, so as to avoid confounding, investigating the role of trust in use of automation therefore requires pairing subjective measurements of the intervening variable, trust, with observations of behavior. To enhance the explanatory power, three types of moderators are included in the proposed model to study how cultural-technological contexts, individual differences, and cultural differences affect trust intention and consequence behaviors, as explained below.

Cultural-Technological Contexts

Cultural-Technological contexts represent the distinct situations of the involved tasks, and pertain to voluntariness of use, risk, complexity, workload, and facilitating conditions.

Individual Differences

Individual differences refer to an individual's background including her age, gender, education, prior experience, instinctive cognitive capacity, and personality traits.

Cultural Differences

Cultural differences contain Hofstede's cultural dimensions (power distance, uncertainty avoidance, masculinity/femininity, and long term orientation).

DISCUSSION

Our initial efforts to develop a psychometrically grounded instrument for measuring trust in automation are described. Eight extant measures were pooled, data collected, and the results subjected to standard scale development techniques. Our results largely confirm the tripartite construct proposed by Lee and See (2004) and largely adopted by researchers investigating trust between people. Items were empirically categorized into two types, general or specific use of automation, and then tested via MTurk. Through EFA, 59 items were extracted to refine the instrument; however, after eliminating the inappropriate items, a total of 42 items were reserved, 26 items falling into general cluster and 16 items involving in specific use of automation. These items were associated with relevant (sub)constructs or moderators, in which 3 main constructs (performance expectancy, process transparency, and purpose influence) and 3 types of moderators (cultural-technological contexts, individual differences, and cultural differences) were involved and framed the proposed model. To further examine the external validity of the model, another round of data collection is being conducted via MTurk.

Existing literature of trust in automation invariably acknowledges the richness and multiplicity of influences. The overall goal of this research is to study both theoretically and empirically the effect of cultural as well as individual contexts on trust antecedents, trust establishment, trust dissolution after the occurrence of faults and trust restoration in human interaction with automation. We expect the results of this research to provide a reliable instrument and a sensitive model that can be used across cultures to measure trust and its antecedents.

ACKNOWLEDGMENT

This research has been sponsored by AFOSR FA9550-13-1-0129.

REFERENCES

Casey, S. M., & Casey, S. (1993). Set Phasers On Stun And Other True Tales Of Design, Technology, And Human Error: And Other True Tales Of Design, Technolo.

Fulmer, C. A., & Gelfand, M. J. (2010). Dynamic Trust Processes: Trust Dissolution and Restoration, 1–32.

Funk, K., Lyall, B., & Wilson, J. (1999). Flight deck automation issues. *Journal of Aviation.*

Goillau, P., & Kelly, C. (2003). Guidelines for trust in future ATM systems: Measures.

Gunia, B. C., Brett, J. M., Nandkeolyar, A. K., & Kamdar, D. (2011). Paying a price: culture, trust, and negotiation consequences. *The Journal of applied psychology*, *96*(4), 774–89. doi:10.1037/a0021986

Halpert, A., Horvath, A., Preston, F., Somerville, K., & Semnani-azad, Z. (2008). *Technological Adoptiveness Scale (TAS): Internal Properties and Construct Validity*.

Hofstede. (1991). Cultures And Organizations - Software of the Mind. *Development*, 1–29.

Hwang, Y., & Lee, K. C. (2012). Investigating the moderating role of uncertainty avoidance cultural values on multidimensional online trust. *Information & management*, *49*(3), 171-176.

Im, I., Hong, S., & Kang, M. S. (2011). An international comparison of technology adoption. *Information & Management*, *48*(1), 1–8. doi:10.1016/j.im.2010.09.001

Jian, J., Bisantz, A., & Drury, C. (2000). Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive*.

Lee, I., Choi, B., Kim, J., & Hong, S. (2007). Culture-technology fit: effects of cultural characteristics on the post-adoption beliefs of mobile internet users. *International Journal of Electronic. 11*(4), 11–51. doi:10.2753/JEC1086-4415110401

Lee, J., & Moray, N. (1992). Trust, control strategies and allocation of function in human-machine systems. *Ergonomics*, *35*(10), 1243–1270.

Madsen, M., & Gregor, S. (2000). Measuring human-computer trust. *11th Australasian Conference on Information Systems*, *53*, 6–8.

Mcknight, D., & Carter, M. (2011). Trust in a specific technology: An investigation of its components and measures. *ACM Transactions on Management Information Systems*, *2*(2). doi:10.1145/1985347.1985353

Merritt. (2000). Culture in the Cockpit: Do Hofstede's Dimensions Replicate? *Journal of Cross-Cultural Psychology*, *31*(3), 283–301. doi:10.1177/0022022100031003001

Meyerson, D., Weick, K., & Kramer, R. (1996). Swift trust and temporary groups. *Trust in organizations.*

Mosier, K. L., Skitka, L. J., Heers, S., & Burdick, M. (1997). Automation bias: decision making and performance in high-tech cockpits. *The International journal of aviation psychology*.

Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. doi:10.1518/001872097778543886