Proceedings of the ASME 2019
International Design Engineering Technical Conferences
and Computers and Information in Engineering Conference
IDETC/CIE 2019
August 18 - 21, 2019, Anaheim, California, USA

IDETC2019-98395

MANIPULATING TRUST OF AUTONOMOUS PRODUCTS WITH AFFECTIVE PRIMING

Ting Liao
Mechanical Engineering
Stanford University
Stanford, California, USA
tingliao@stanford.edu

Erin MacDonald
Mechanical Engineering
Stanford University
Stanford, California, USA
erinmacd@stanford.edu

ABSTRACT

User-to-product trust has two notable aspects: (1) the user's propensity to trust, and (2) the product's trustworthiness as assessed by the user. Autonomous products, which perform many functions on their own with limited user input, require the user to exhibit trust at an appropriate level before use. Research in product trust thus far has focused on the product trustworthiness: manipulating the product's design, for example, anthropomorphizing an autonomous vehicle and measuring changes in trust. This study flips the usual approach, manipulating a person's propensity to trust and measuring response to an existing autonomous product, the Amazon Echo. We build on our past successes with priming exercises to reveal insights into the user-related factors of product trust. In this study, we used visual stimuli that evoked either positive, neutral or negative emotions as affective primes to influence users' trust propensity before the interaction. The participants interacted with a mock-up of the Amazon Echo via ten pre-determined question-and-answer (Q&A) sets. During the interaction, the participants evaluated the Echo's competence and if it met participants' expectations. They also reported trust towards the Echo after the Q&A sets. Holistically, the affective primes show no significant effect on the trust propensity. For the subgroup of participants whose expectations of the product's performance were met, both the perceived product competence and the affective primes have significant effects on trust propensity. These results demonstrate the complex nature of trust as a multidimensional construct and the critical role of product performance in trust formation. They also suggest that it will be difficult for a product to build trust with users who expect the product to perform in a different way than its intent—*if one wants to design a product that builds trust, they should understand user expectations and design to meet them.* This learning can facilitate the intentional design of the affective process in trust formation that helps build a healthy level of trust with autonomous products.

NOMENCLATURE

T_{ij}	Trust-related belief about the system of j's participant measured by i's question in post-interaction survey
C_j	Test condition assigned to participant j
$\overline{PPC_j}$	Average perceived product competence across all ten questions of participant j
A_{kj}	Initial attitude towards machines in general of j's participant measured by k's question in pre-interaction survey
$PPC_{Q_x,j}$	Perceived product competence for the interactive question x of participant j
$\beta, \varphi, \omega, \gamma, \varepsilon$	Estimated coefficients of the regression models

1. INTRODUCTION

In the modern world, many automated systems, such as GPS navigation systems, are used to assist people in acquiring information, making decisions, and carrying out actions for various tasks [1]. While automated products and systems can drastically improve the efficiency of labor systems, grave consequences can occur if users over-rely on the automated systems or underutilize them—both of which can occur due to an inappropriate level of trust [2]. Like its role in an interpersonal relationship, trust is a critical factor in determining the willingness of users to adopt and rely on automated systems in situations with uncertainty [1].

Social interaction and interpersonal relationships require trust, and thus, researchers have studied it among individuals and organizations. Although definitions of trust vary across the fields of research, commonality exists between interpersonal trust and product trust. At the fundamental level, both types of trust represent context-specific attitudes in situations characterized by uncertainty [1]. Trusting is a dynamic process; it evolves over time and across interactions as the continuous assessment of trustworthiness feeds back into the ongoing trust appraisal of the person disposing trust [3].

As people behave in interpersonal relationship, research by Nass and Moon showed that people applied social rules when interacting with inanimate machines [4]. Damen and Toh further explored the impact of social rules, such as gender stereotype, on trust-related behavior and validated that users extended their social expectations to their interactions with home automation simulations [5]. When products become more interactive and autonomous, the social interactions and user expectations of these products are critical to explore.

Lee and See described trust as a metal state that involves both thinking and feeling when trusting automated products [6]. Feeling, also referred to as emotion or affect, is an important but overlooked part of trusting automated products.

In this work, we intentionally alter affect. One of the robust methods to alter a user's affect is priming via exposure to a stimulus that activates an idea, contextualization, or feeling [7]. To investigate the social interactions between users and products with a focus on the affective process of trust formation, we proposed an approach to subconsciously manipulate user-to-product trust by priming users with emotions prior to the interaction, potentially making users feel slightly happy or sad. We tested to see if priming positive or negative emotions can effectively influence users' mental state regardless of the performance of the product and consequently increase or decrease participants' trust propensity.

As we aimed to focus on social interactions and affective process of trust formation in products, we chose a mock-up version of the Amazon Echo as the focal product. The Amazon Echo is a smart speaker (conversational agent) that connects to cloud service to play music, answer questions, etc. [8]. It incorporates a humanized voice and enables autonomous responses and simple social interactions.

In this study, we randomly assigned participants into three test conditions: (A) prime with positive emotions, (B) prime with neutral emotions, and (C) prime with negative emotions. All participants, working alone, reported their initial attitude towards machines, and then were primed with a set of visual stimuli embedded with either positive, neutral, or negative emotions. Afterwards, they interacted with a focal product, a mocked-up version of the Amazon Echo, with pre-determined answers, and evaluated its performance. After the interaction, the participants self-reported how much they trusted the product. We tested if priming with positive emotions would encourage users to trust the product more compared to priming with neutral emotions, and vice versa for negative emotions.

This paper is organized as follows: Section 2 provides a review of user-to-product trust formation, and methods to prime trust and measure trust in technology-related fields. Section 3 lists our proposition and hypotheses. Section 4 gives an overview of the experiment procedure, and Section 5 provides more details about how we prepared the experiment. Data and analysis are in Section 6, and discussion of the results in Section 7. Finally, Section 8 includes the conclusion and plan of future work.

2. BACKGROUND

2.1 User-to-Product Trust Formation

Trust was first studied within the context of interpersonal relationships and rooted in individual willingness-to-accept vulnerability [9]. As more human jobs are replaced by automated systems, the theory of interpersonal trust has been expanded to user-to-product relationships, and in particular, the concept of *trust in automation* has become an important area of research in recent years [1]. Rather than defining trust as an action item as defined in interpersonal domain, researchers in the field of trust in automation view trust as a mental state and adopt the definition of trust by Lee and See as "the attitude that an agent will help achieve an individual's goal in a situation characterized by uncertainty and vulnerability" [6]—it is a task-related and context-dependent property.

Regardless of its domains, almost every explanation of trust involves a trustor to give trust, a trustee to perform the task and accept trust, and something at stake – there must by a possibility that the trustee will fail to perform the task, inviting uncertainty and risk. Characteristics of both trustor and trustee as well as context interplay to form trust. The "Big Three" predictors in the classic trustworthiness research by Mayer et al. in the interpersonal relationship are *ability* (perceptions of a trustee's competence and consistency), *benevolence* (perceptions of the trustee's caring, goodwill, empathy, and commitment to shared goals), and *integrity* (perceptions of the trustee's objectivity, fairness, honesty, and dedication) [9]. In user-product relationships, the predictors to describe trustworthiness are reduced to *ability* or *competence*. They are further categorized as the performance-based components including predictability, dependability and reliability, and the attribute-based components, such as product type (e.g. voice agent, etc.) [10]. For trustors, human-related factors such as gender, personal traits, and ages play important roles in the human-product relationships as well [1].

Trust researchers in the areas of Human-Computer Interaction (HCI) and Human-Robot Interaction (HRI) extensively studied the performance-based components and competence of the system itself, for example, level of automation and information transparency. Hancock et al. [10] conducted a meta-analysis and tested the correlations between the influencing factors and trust in HRI. They suggested that robot characteristics, and in particular, performance-based factors had the largest current influence on the perceived trust in HRI; they also mentioned that human-related and environmental factors were not examined in the meta-analysis due to insufficient samples from the existing research [10].

While the human-related factors are sometimes overlooked, the social aspects are inevitable in the user-product relationships, especially when the systems become more autonomous. Nass et al. argued that users apply social rules to computers, even though such attributions may not be needed for inanimate objects [11].

Imbuing such social cues in the interaction has the potential to increase trust of autonomous products as reflected in both behavior and self-reported measures [12] and an effective approach is anthropomorphism. Designers implemented anthropomorphism by embedding inanimate products with human-like features, for example, facial expression, voice tone, and personality [13]. Lopatovska and Williams [14] found that users reported higher satisfactions if they named their home automation product as a person. This finding indicates that how humans interact with technology towards a product can vary depending on how the product presents itself. Despite the advantage of imbuing social aspects however, Culley and Madhavan brought up a caution that the emotional connection could lead to over reliance on the autonomous systems and inappropriate trust calibration [15].

The research mentioned above highlights the importance of the social cues in user-product relationship and also the complex nature of user trust within social context.

2.2 Priming Trust

Like all other decisions, trust is driven by a combination of cognitive (thinking) and affective (feeling) factors. Mental state is part of social interactions and the premise for trust decisions. Priming is a robust psychological technique that has been extensively used in psychology, behavioral economics, and organizational behavior to activate specific mindsets or mental states. Researchers have used it to inspire sustainability concerns [16], facilitate designers to communicate sustainability [17] and generate environmental-friendly ideas [18].

The psychological mechanism of priming is developed from the concept of "perceptual readiness" proposed by Bruner [19]. Bruner claimed that the information and feelings that are currently cognitively accessible lead to corresponding thoughts and behavior [19]. Priming activates a specific set of information and feeling by increasing the accessibility of relevant thoughts, memories and feelings and consequently, it motivates related perspectives, decisions, and behavior [20]. Previous research by She and MacDonald has shown promise to activate mindsets for design tasks via a collage priming method, where participants physically arrange product images to create a paper collage [10, 11]. She and MacDonald found the method to be effective in establishing product semantics for sustainable products and in inspiring sustainability concerns [17]. Liao and MacDonald further applied the collage priming method to facilitate designers in improving ideas related to environmental impact in eyes of users and to encourage a more receptive attitude towards hypothetical others' ideas in a group setting [18]. In addition to the priming methods with a focus on cognitive process, Lewis et al. conducted a study that utilized the affective priming to increase the quality of design ideas generated by showing participants a picture of a laughing baby [21].

In addition to the applications in design tasks, research has also applied various priming methods to specifically change trust. A study by Al-Ubaydli showed that participants being exposed to trade-related concepts expressed higher level of trust towards strangers compared to ones being exposed to non-trading words. The authors argued that market proliferations allowed good things to happen when interacting with strangers, thus encouraging optimism and trusting behavior [22]. In contrast of trade-related terms, after being primed with legal concepts, for example, terms used in lawsuits, the participants perceived social actors as less trustworthy and the situation as more competitive [23]. Besides text-based stimuli, visual images that evoke a particular emotion have been shown to influence trust. Brownlow [24] found that baby-faced speakers induced more agreement with their position than mature-faced speakers did when trustworthiness was in question, presumably because baby-faced speakers still appeared honest due to their babyish facial features.

Besides associating with particular content, visual stimuli that trigger designated emotions can influence trust. A study by Hooker et al. showed that exposure to negative (threat-related, e.g. weapons) images led people to rate unfamiliar faces as less trustworthy compared to priming with neutral or positive (e.g. puffy animals) affective pictures [24]. However, the affective visual stimuli have the contrast effect in the interactions with automation. In a study by Rice et al., priming users with negative images, such as beat-up cars, prior to the interactions with a recommendation agent increased reliance on such agent in a combat identification task, while decreasing the reaction time and increasing accuracy for the task, compared to priming with positive images of luxury cars [25].

To date, numerous studies have validated the effect of affective images on both self-reported trust and trust-related behavior, whereas the dissonance of self-reported and behavioral measures of trust exists. In this study, we implement affective images with normative valence scores from an open-source database to evoke emotions and measure the change in trust of the focal autonomous product. The selection of affective images is described with details in Section 5.1.1.

2.3 Measuring Trust

When investigating a complex construct like trust, one of the main challenges is to identify a robust method to appropriately measure trust between users and products. Previous research has explored both implicit and explicit methods, where explicit methods tend to capture conscious and immediately observable behavior and implicit methods measure unconscious actions, such as respiration rate, heart rate, and reaction time [26]. While both explicit and implicit methods have their own merits, different methods can lead to different results [27] and this dissonance highlights the complex nature of understanding trust and the lack of full awareness of people's own inner thoughts, emotions and behavior.

Even though researchers have utilized implicit measures such as reaction time to reflect unconscious attitude and quantify trust-related behavior of the automated products, their causal relationship with trust in technology has not been fully validated. As there exists insufficient evidence of the correlation between explicit and implicit measures, this study will mainly use explicit metrics adopted from the research of the trust in interpersonal relationship and self-reported data.

Researchers have invented various questions that best suited their purposes and context to measure trust, even though the core of what they wanted to know is identical—how much users trust the product. Yeh and Wickens [28] measured trust in the design of augmented reality by directly asking "How much did you trust the terrain information presented in the computer-generated imagery?"

Besides simply asking about trust, Koo et al. [29] associated "*machine acceptance*" with "*emotional valence*" in a post-drive questionnaire to assess drivers' attitudes. The participants ranked the provided adjectives, such as "anxious," "annoyed," and "frustrated", on a ten-point Likert scale ranging from "Describes Very Poorly (=1)" to "Describes Very Well (=10)" in response to the question, "How well do the following words describe how you felt while driving". The machine acceptance index reflects responses to the question, "How well do the following adjectives describe the car" [29]. However, in this study, the correlation between emotional valence and machine acceptance was not supported.

Another common approach is to break down trust into relevant determinants that tend to be more tangible. Muir and Moray [30] captured operators' trust in a process control simulation by a set of subjective rating scales for the pump subsystem in terms of its *competence* ("to what extent does the pump perform its function properly"), *predictability* ("to what extent can the pump's behavior be predicted from moment to moment"), *dependability* ("to what extent can you count on the pump to do its job"), *responsibility* ("to what extent does the pump perform the task it was designed to do in the system"), and *reliability over time* ("to what extent does the pump respond similarly to similar circumstances at different points in time").

Building upon the previous work, Jian et al. conducted an empirical study with correlation and cluster analysis of the trust-related concepts that were widely used [31]. They identified the similarities and differences among those concepts of generalized trust, trust between humans, and trust between humans and automated systems. The study provided a model for assessing trust between humans and automated system based on empirical data [31].

Many explicit metrics provided in this section were tailored to their focal products and context. To cover a broad spectrum of potential determinants of trust, such as security and reliability, and make sure the questions to remain interpretable, we choose to adopt the questions from the original work by Jian et al. [31] that has been validated on the applications of smart systems in cars [32] and personal remote assistant [33]. The list of questions is in Section 5.2.

3. PROPOSITION AND HYPOTHESES

Proposition: *A positive prime on emotional affect (making someone feel slightly happy) can increase users' propensity to trust a product compared to the neutral prime; a negative prime (making someone feel slightly sad) can decrease users' propensity to trust the product compared to the neutral affective prime.* Priming emotional affects has the potential to change people's subconscious attitudes and we hypothesize that people being exposed to positive emotions will also exhibit a more positive attitude towards the product and accordingly higher trust level compared to the ones being exposed to the neutral emotions. To determine if affective priming successfully alters participants' attitudes towards the product, we compare the self-reported trust level towards the focal product after the interaction between either the positive or negative prime and neutral prime.

Hypothesis 1: Priming with positive emotions results in higher level of self-reported trust towards the focal autonomous product than priming to neutral emotions.

Hypothesis 2: Priming with negative emotions results in lower level of self-reported trust level towards the focal autonomous product than priming to neutral emotions.

4. EXPERIMENT OVERVIEW

In this section we give an overview of the experiment, while more details about the experiment design are provided in Section 5. We had three test conditions: A) prime with positive emotions, B) prime with neutral emotions, and C) prime with negative emotions. Figure 1 illustrates the four major steps in the experiment.

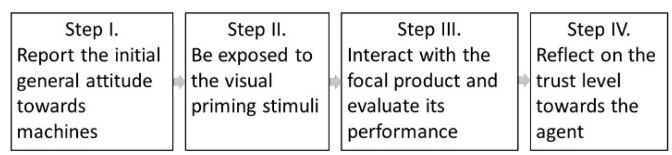


FIGURE 1: OVERVIEW OF EXPERIMENT PROCEDURE

The study was conducted in a room that contained a laptop for answering questions and the focal autonomous product, a mock-up of the Amazon Echo (Fig. 2). Participants were prompted to ask the Echo a set of questions in a random order in Step III. A proctor sat in an adjacent room that could not be seen by the participants (Fig. 3) but the proctor could hear sounds from the experimental room via a microphone. The proctor heard the questions asked by the participants and played the corresponding oral responses through a Bluetooth speaker that was hidden under the focal product. The oral responses were recorded and selected from the real responses of the Amazon Echo. The recorded responses were used to control the experiment. The participants were deceived about the fact that the responses were not delivered by the Amazon Echo in real time. All participants performed the activity alone.

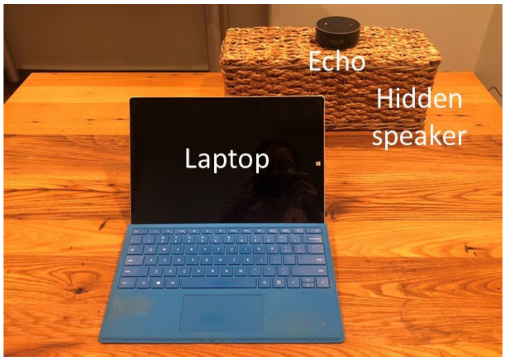


FIGURE 2: ILLUSTRATION OF THE SETUP IN EXPERIMENTAL ROOM

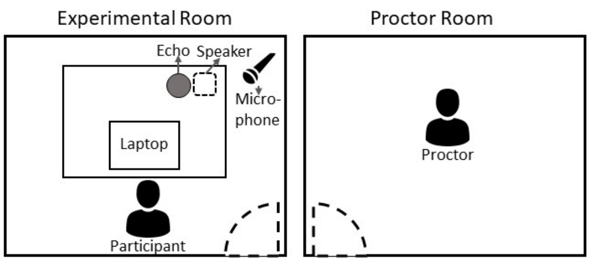


FIGURE 3: DIAGRAM OF THE FULL SETUP

The experiment procedure consists four steps as described in Fig. 1:

I. Pre-test Survey: Participants take a survey about their demographic background and general attitude towards machines.

II. Priming: Participants are exposed to a set of visual stimuli that evoke either positive, neutral or negative emotions according to their assigned test conditions.

III. Interaction: Participants are prompted to ask the mock-up of the Amazon Echo ten questions in sequence of a random order and listen to its oral responses. After each round of question and response, participants rate the perceived competence of each response and report if the product performance meets their expectation.

IV. Post-test Survey: Participants report their trust propensity of the product that they interact with in Step III via a set of trust-related metrics.

5. EXPERIMENT DESIGN

In this study, we had three test conditions, and in each condition, we used affective images as visual stimuli to create a particular emotional affect (slightly happy or sad or neutral) to the participants. Participants interacted with the focal product via a set of the pre-determined questions and responses that represented various levels of performance of the product. The selections of the visual stimuli and the questions prompted to the participants along with the corresponding responses are explained in Section 5.1.

There are many autonomous or "smart" products existing in the market, and here we chose the Amazon Echo as the focal autonomous product for the following reasons: 1) a voice recommendation agent features simple oral conversations and its performance can be easily justified by the participants; 2) Echo has been widely available in the U.S. since 2015 [6] and we expected that participants would have at least heard of it and would not feel awkward talking to this inanimate object. The Amazon Echo Dot is a smaller version of the Amazon Echo with the same capabilities. It is a hockey puck-sized product (Fig. 2) and it was placed in front of the participants in the study.

In the experiment, we measured participants' initial attitude towards machines in Step I and perceived product competence and expectation (satisfaction) of each response in Step III. After the interaction, we collected trust-related measures in Step IV. The lists of questions to assess the trust-related attitude prior and post to interaction as are described in Section 5.2 with details.

5.1 Stimuli Selection

5.1.1 Visual Stimuli

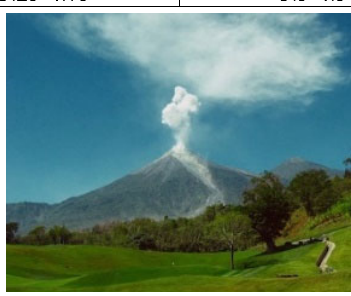
The visual stimuli were selected from the Open Affective Standardized Image Set (OASIS), an online database that contains 900 color images with normative ratings on two affective dimensions-valence (i.e., the degree of positive or negative affective response that image evokes) and arousal (i.e., the intensity of the affective response that image evokes) [34]. These images depict a broad spectrum of themes, including humans, animals, objects, and scenes, and they were rated using a 7-point scale by a sample of participants in the U.S. recruited via Amazon Mechanical Turk [34]. Before OASIS was constructed, the International Affective Picture Systems (IAPS) by the Center for the Study of Emotion and Attention has been widely used in psychological and neuroscience research [35]. However, the original version of IAPS was produced in pre-Internet time and the scores were assigned by a group of college students in 1997 [35]. To incorporate the latest images and rating scores by a more representative population, we decided to choose our visual stimuli from OASIS.

We selected 54 pictures, including 36 evocative (18 positive and 18 negative) and 18 neutral ones based on their content and normative scores of valence (hedonic state, e.g. positive emotions and negative emotions) and arousal (activation). In this study, as we investigated the effect of the degree of positive or negative affective responses, we only varied the valence score, while the score of arousals is kept at the neutral level (score = 4 ± 0.5) to avoid unnecessary excitement during the experiment. The images were not necessarily associated with a particular emotion, yet the emotion with a positive valence often refers to happiness. Pictures that trigger anger were likely to be excluded for this study because we selected pictures with relatively neutral arousal, even though anger and rage sometimes occur when users are disappointed by product's performance. The desired ranges of the valence scores are listed in Table 1 for positive, neutral and negative primes, along with an example of each condition.

Pictures with extreme scores were intentionally left out, for example, the ones with valence scores less than 0.25 or greater than 6.75. Regions between two groups, for example, 2.75–3.25, were left out to avoid ambiguity. Pictures showing inappropriate content, such as blood, scars, wounds, disease, weapons, war, death, religion, nudity, and eroticism/romance were excluded as well. Images that directly related to machines were eliminated to avoid conscious bias and fixation.

Regarding stimuli presentation, a common method to reinforce memory in cognitive and psychological research is a combination of massed repetition (i.e., repeating consecutively) and distributed repetition (i.e., repeating over a longer period of time). Massed repetition has shown to result in a better memory performance at short retention intervals, and distributed repetition has shown to be effective in reinforcing the long-run memory [37, 38].

TABLE 1: DESIRED RANGES OF VALENCE AND AROUSALS SCORES FOR VISUAL STIMULI

Positive	Range of Valence	Range of Arousal
	5.25-6.75	3.5-4.5
		
Neutral	Range of Valance	Range of Arousal
	3.25-4.75	3.5-4.5
		
Negative	Range of Valance	Range of Arousal
	0.25-2.75	3.5-4.5
		

For each condition, a participant was presented 18 images, of which 6 pictures were presented only once, 6 were repeated three more times consecutively (18 images massed repetition), and 6 were repeated three more times distributed across the study (18 trials, distributed repetition). Each participant saw 54 images in total and each picture was displayed for 4 seconds.

5.1.2 Interactive Questions

In Step III, participants interacted with the mock-up of the Amazon Echo by asking it ten pre-determined questions and hearing its responses (see Table 2). To control the experiment, we asked the participants to read out the exact questions prompted on the laptop screen. The corresponding oral responses were recorded from the Amazon Echo ahead of time and played via a Bluetooth speaker that was hidden under the mock-up Echo (Fig. 2). The mock-up Echo was presented in the study room and the participants were deceived about the fact that they were not interacting with a fully functioning product. Regardless of how participants asked questions (with pauses, accents, or mumbling, etc.), as long as the full question prompted to them was read out, a response would be delivered.

TABLE 2: QUESTIONS PROMPTED TO PARTICIPANTS DURING INTERACTION AND CORRESPONDING ANSWERS

Index	Script
Q1	How do I make a pumpkin spice latte?
A1	Okay, for pumpkin spice latte, I recommend pumpkin spice latte, which takes 7 minutes to make.
Q2	How do I get to Stanford Shopping Center?
A2	The fastest route to Stanford Shopping Center at 855 El Camino Real takes about 10 minutes via Campus Drive and Quarry Road.
Q3	What food is toxic to dogs?
A3	A duck's diet is wide and varied. It includes small fishes and crustaceans, fish eggs, snails, worms, mollusks, grass, wheats, algae, aquatic plants and roots, frogs, salamanders, seeds, grains, berries, fruits, and nuts.
Q4	When was Stanford founded?
A4	Here is something I found in Wikipedia. Stanford University is a private research university in Stanford, California, in Silicon Valley, 20 miles outside San Jose.
Q5	When is the Arrillaga Outdoor Education and Recreation Center (AOERC) open today?
A5	The Arrillaga Outdoor Education and Recreation Center on Santa Teresa street is open now till 12 AM.
Q6	What day is Memorial Day 2018?
A6	In 2018 the Memorial Day is May 28th, 2018.
Q7	How long is a typical flight from San Francisco to Tokyo, Japan?
A7	The flight time between San Francisco and Tokyo prefecture is approximately 11 hours.
Q8	How do I use an electrical press drill?
A8	Sorry I don't know that one.
Q9	Who invented mobile phones?
A9	Mobile phone inventors are AT&T and Martin Cooper.
Q10	How much protein is in two eggs?
A10	Here is what I found in the internet: two raw eggs contain 12 grams of protein.

To determine the stimulus questions and responses, we conducted a pilot study in which a few graduate students at Stanford University asked Echo a number of questions that they would have asked in daily life. The questions and oral responses were selected to represent various levels of accuracy, usefulness and ease of interpretation. Non-informative (e.g. A1 and A2) irrelevant responses (A3) and error message (A8) were selected to mimic the realistic performance of the Echo and to introduce uncertainty, which is a premise for a trust decision.

5.2 Metrics

As mentioned in Section 2.3, trust is a dynamic state and multidimensional construct. Researchers have invented both explicit and implicit metrics to assess it. In this study, we used directly reported measures of trust prior- and post-interaction. In addition, as we anticipated that the most competent response computed by computer algorithms might not be perceived as the most useful or satisfactory by users, we measured *user-perceived* product competence and user expectation right after each oral response was delivered by the mocked up Echo, the participants rated the perceived product competence by answering the question "*does this response answer your question about …*" on a 5-point scale, and reported their expectation level by answering the question "*does this response meet your expectation*".

The questions to assess the trust of the product and its potential subconstructs (e.g. dependability, reliability) were adopted from the original work by Jian et al. [31] that measure trustworthiness of automation technology. The questions were tailored to best suit the focal product and context of this study. The participants reported if they agreed or disagreed with the statements in Table 3 for the initial attitude towards machines and with the statements in Table 4 for their trust-related beliefs of the system on a 5-point scale (1 = totally disagree, 5 = totally agree) after the interaction.

TABLE 3: PRE-INTERACTION SURVEY QUESTIONS ABOUT INITIAL ATTITUDES TOWARDS MACHINES

Index	Statements
1	For the most part, I distrust machines.
2	I am likely to trust a machine even when I have little knowledge about it.
3	I usually trust machines until there is a reason not to.
4	In general, I would rely on a machine to assist me.
5	It is easy for me to trust machines to do their job.
6	My tendency to trust machines is high.

TABLE 4: POST-INTERACTION SURVEY QUESTIONS ABOUT TRUST-RELATED BELIEFS OF THE SYSTEM

Index	Statements
1	The system is deceptive.
2	I am suspicious of the system's intent, action, and output.
3	I am wary of the system.
4	The system's actions will have a harmful or injurious outcome.
5	I am confident in the system.
6	The system provides security.
7	The system is dependable.
8	The system is reliable.
9	I can trust the system.
10	I am familiar with the system.

6. DATA AND ANALYSIS

Forty-eight people participated in the study and each was compensated with 15 U.S. dollars for their time. As the interaction with Echo was mimicked and not actual, some features like glowing when being called were disabled. Four respondents reported a malfunction of the focal product by noticing things like the Echo not glowing and interrupted the experiment. Their responses were excluded in the analysis.

The following analysis includes forty-four valid responses. The number of participants in each condition is shown in Table 5. Forty participants are Stanford affiliated. Of the total number of participants, thirty-two are female, and twelve are male.

TABLE 5: NUMBER OF PARTICIPANTS IN EACH CONDITION

Condition	Negative	Neutral	Positive	Total
# Participants	14	17	13	44

6.1 ANOVA

Analysis of Variance (ANOVA) is used to test if there is significant difference between group means of three test conditions. This test is applied to all the raw measurements, including the initial general attitude, perceived competence of each oral response, expectation, and post-interaction attitude as measured by the level of agreements with the statements shown in Table 4. ANOVA did not indicate a group difference on the initial general attitude towards machines. This indicates that participants had a relatively homogenous attitude towards machines

6.2 Linear Mixed Model with Linear Terms

A linear mixed model (LMM) described in Eqn. (1) is used to control the effect of individual differences for the initial attitudes towards machines and test the effects of the priming condition and the perceived product competence of the responses.

$$T_{ij} = C_j \cdot \beta_i + \overline{PPC}_j \cdot \varphi_i + \sum_{k=3,5,6} A_{kj} \cdot \gamma_{ijk} + \varepsilon_{ij} \qquad (1)$$

Where C (testing condition) and $\overline{\text{PPC}}$ (the average perceived product competence across all ten Q&A sets) are considered as the fixed effects and A (initial general attitude) is the random effect; i represents the index of the trust-related metrics (Table 4), j represents the individual participant, and k represents the index of the questions for the initial general attitude (Table 3).

The results show that the positive prime has a significant effect (est. coefficient = 0.501, p = 0.027*) on the self-reported agreement about the statement "*the system's actions will have a harmful or injurious outcome*". The average perceived competence of all responses has a significant effect (est. coefficient = -0.314, p = 0.006*) on the statement "*the system's actions will have a harmful or injurious outcome*". This result indicates that participants being exposed to the positive priming stimuli were more suspicious about the system's outcome, and the performance is an essential factor to justify the outcome of the system.

The positive prime also shows the significant effect (est. coefficient = 0.671, p = 0.023*) on agreement with the statement "*the system provides security*" (Fig. 4). The participants being exposed to the positive priming stimuli considered the system to be more secure compared to the participants being exposed to the neutral stimuli. Thus, with the positive prime, participants found the system more secure than under other primes, but they did not trust the responses more or expected better product performance.

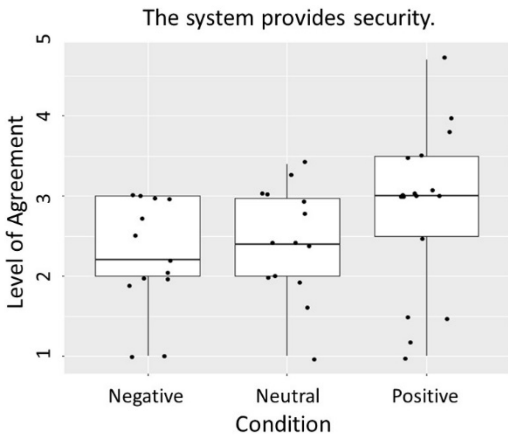


FIGURE 4: SELF-REPORTED AGREEMENT ON THE STATEMENT "THE SYSTEM PROVIDES SECURITY."

With LMM shown in Eqn. (1), the priming conditions and average perceived competence are tested to have no significant effect on participants' agreement on the statement "*I can trust the system*".

6.3 LMM with Median Separation and Interaction

As shown in Section 6.2, the perceived product competence can influence participants' agreement on the statements about product's performance and further outcome. To further explore this, data is split at the median value (median = 3.40) of the average expectation scores for each participant across all ten stimulus questions (Table 2) measured by answering the question "*does this response meet your expectation*". The median value is used instead of the absolute neutral value (= 3) to compensate for the tendency to report high expectation values. We focus on the subset of data with the expectation scores above the median in this part. An LMM model as described in Eqn. (2) is again used, with the addition of interaction between the priming condition and the perceived product competence of the responses to Q1 ("*how do I make a pumpkin spice latte*") and Q2 ("*how can I get to Stanford Shopping Center*").

$$T_{ij} = C_j \cdot \beta_i + PPC_{Q1,j} \cdot \varphi_i + C_j \cdot \omega_i \cdot PPC_{Q2,j} + \sum_{k=3,5,6} A_{jk} \cdot \gamma_{ijk} + \varepsilon_{ij} \tag{2}$$

Where similarly, β, φ, ω represent fixed effects parameters and γ represents the parameter of the random effect.

The perceived product competence of the responses to Q1 and Q2 are incorporated in this model because these two responses (shown in Table 2) were reported as being difficult to justify as their perceived product competence scores show large variances (Fig. 5). The variance of Q1 competence is 1.117 and the variance of Q2 competence is 1.427. As mentioned in Section 1, the presence of uncertainty involves in the trust formation, and the effect of those *uncertain-to-justify* responses on product trust is interesting to explore. In contrast, for example, the perceived competence of responses to Q7 ("*how long is a typical flight from San Francisco to Tokyo, Japan*") and Q8 ("*how do I use an electrical drill press*") have relatively small variances (variance (Q7) = 0.250 and variance (Q8) = 0.028) as shown in Fig. 6.

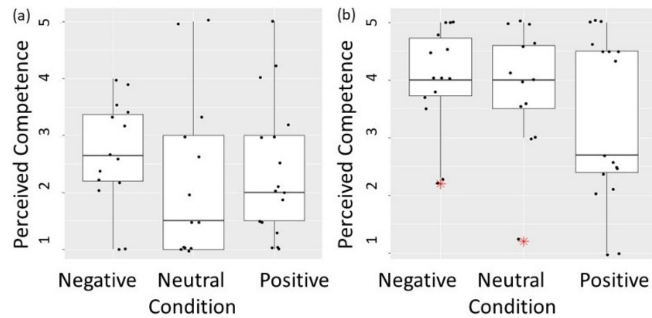


FIGURE 5: BOX PLOTS OF PERCEIVED PRODUCT COMPETENCE OF (a) Q1 AND (b) Q2

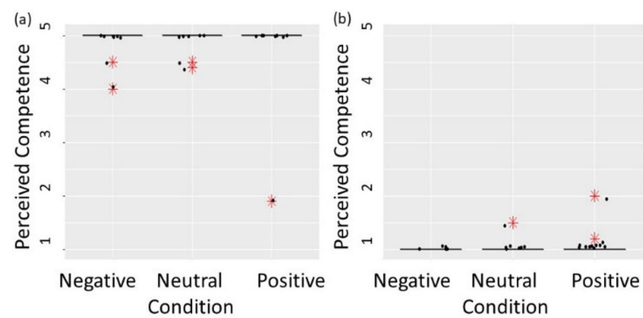


FIGURE 6: BOX PLOTS OF PERCEIVED PRODUCT COMPETENCE OF (A) Q7 AND (B) Q8

The results show that the perceived product competence of Q1's response (abbreviated as Q1 competence) has a significant effect (est. = 0.491, p = 9.90e-6*) on participants' agreement on the statement "*I can trust the system*", and the interaction between the negative prime and Q1 competence also has a significant effect on this measure (est. = -0.520, p = 0.010*). Q1 competence also has a significant effect (est. coefficient = -0.245, p = 0.003*) on participants' agreement on the statement about the system's harmful outcome, which is consistent with the result obtained in Section 6.2.

In addition, both the perceived product competence of Q2's response (Q2 competence) and the interaction between primes and Q2 competence have significant effects on the trust level. The estimated coefficients and corresponding p-values are shown in Table 6. The results indicate that both positive and negative primes enable changing participants' agreement on the statement "*I can trust the system*", and so does the higher perceived product competence.

TABLE 6 ESTIMATED COEFFICIENTS OF Q2 COMPETENCE AND PRIMES ON AGREEMENT ON TRUST STATEMENT

Fixed Effects	Estimate	P-value
Negative Prime	-6.22	0.004*
Positive Prime	-7.97	0.013*
Q2 Competence	-1.30	0.011*
Q2 Competence x Negative Prime	1.54	0.010*

7. DISCUSSION

The experiment demonstrates that priming participants with visual affects is partially effective in changing users' propensity to trust an autonomous product, with the understanding that the perceived quality of the product's performance is a critical factor in this relationship, and user expectation can be a key factor as well.

The hypotheses H1 and H2 are rejected as primes are tested to have no significant effect on changing trust propensity holistically. As mentioned, trust is a complex construct and the effect of the affective primes cannot be simply described by a simple linear model.

However, while the primes do not directly change participants' propensity to trust the product, the positive prime has a significant effect on the self-reported agreements with the trust-related statements, such as "*the system's actions will have a harmful or injurious outcome*" and "*the system provides security*". The results show that exposing participants to the positive priming stimuli created suspicions about the system's outcome, but they were also more likely to agree that the system provided security. Although Jian et al. considered the measures about system's reliance, security, etc. (shown in Table 4) to be aligned with the measure of trust, the results show that they do not necessarily link in new context, at least for the measures of system's security and its harmful outcome. Users may interpret security in a different way, for example, as information security when the responses are produced by a cloud service.

This also indicates that a holistic measurement of trust is not the best approach for measuring subtle changes in trust, potentially due to performance flaws in autonomous products. People are probably more likely to relate flaws to more tangible trust constructs, such as outcomes (audio responses and how they are used, in this study) than intangible relationship factors.

When only focusing on the dataset by participants whose expectations of product performance were met, both perceived product competence and the interaction between the primes and competence have significant effects on self-reported trust. The perceived competence of the *uncertain-to-justify* response to stimulus question Q1 has a significant positive effect on participants' agreement on the statement "*I can trust the system*", and the interaction between the negative prime and Q1 competence has a significant negative effect on this measure.

In addition, surprisingly, both primes and the perceived product competence of Q2 that the Echo gave an *uncertain-to-justify* response for have significant negative effects on self-reported trust, whereas the interaction between negative prime and the Q2 competence has a significant positive effect on the trust. This result suggests that participants being exposed to negative affective stimuli and encountering uncertain responses express more trust compared to neutral affective stimuli. We notice that though Q1 and Q2 are considered to have *uncertain-to-justify* responses, their impacts on product trust contradict to each other. This result is probably due to an underlying threshold of "satisfied" performance of the product, as the Q1 competence with the mean value of 2.18 on a 5-point scale has the opposite effect of the Q2 competence with the mean of 3.71 (while 3 being the neutral point). This suggests different levels of competence, and that our binary classification system of certain/uncertain answers could use further explanation.

This result shows that for participants whose expectations of the product's performance are met, the user perceived product competence plays a leading role in the trust formation—if the expected performance matches the reality it may form a basis for trust. The responses that participants found difficult to interpret or justify have strong influence on trust formation.

This finding supports the design recommendations by MacDonald and She for communicating sustainability to users [11, 38], which suggest that people will trust sustainable products only if the products include the features that the user thinks are sustainable—matching user expectations to build trust is the key.

The above represents the most important takeaway for autonomous product designers: if one wants to build trust, understand user expectations and meet them. Then, work to design the bells and whistles of trust, such as building an emotional connection.

There are several limitations of this study. First, the type of interaction in this study was solely oral and behavioral data were not collected or analyzed in this study. In addition, only self-reported data on trust is used to test the hypotheses. Even though a previous study by Merritt et al. [39] showed a disagreement between the explicit and implicit measures of trust when participants perceived automation, an implicit measure, such as reaction time, can supplement the self-reported survey and enrich study findings. Moreover, the Amazon Echo Dot is a product available in the market, and previous experience interacting with it can bias or distort the results. Additionally, the participants were either students at Stanford University or Stanford affiliates. We had a limited access to the participant population due to the physical location where this study took place, and a more diverse demographic would be ideal. Another limitation we acknowledge is that we used visual stimuli to create emotional affects and aimed to subconsciously change users' trust propensity; however, as the whole priming process intended to be subconscious, we were unable to validate if the designated emotions were effectively evoked, although previous work suggests this is the case. We used visual prime and audio interaction and aligning the channels by using audio prime can possibly produce stronger priming effect for audio interaction. Last, we formulated the questions that participants asked the Echo based on the hypotheses at hand, and the final set of questions were selected by a small group in a pilot study—the basis from the researchers and the pilot participants can potentially influence the results.

8. CONCLUSION AND FUTURE WORK

This research provides further evidence of the difficulties of studying trust in social interactions between users and products. In this study, we primed the participants with visual stimuli that evoked either positive, neutral or negative emotions before their interaction with the focal autonomous product. We measured the participants' perceived competence during the interaction and their self-reported trust towards the mock-up version of the Amazon Echo afterwards.

The results suggest that some components, such as security, that have been considered the main determinants of trust may no longer be good measures of trust in a highly interactive environment, as participants may interpret these components in a different way (e.g. security as information security instead of product security). The study also indicates that a holistic measurement of trust may not be effective to measure the context-specific trust in the interactions with autonomous products.

When only considering the participants, whose expectations were met by the Echo's responses, the perceived competence of *uncertain-to-justify* responses have a significant effect on the participants' trust towards the product, but the direction of the effects for these two *uncertain-to-justify* Echo stimuli contradict. The results also demonstrate a significant effect of the interaction between the perceived product competence and primes on trust propensity. However, the effects of the primes and perceived competence on trust are found to only exist for participants whose expectations were met. In other words, it is even more difficult to predict or construct trust for users with unsatisfied expectations. Thus, this study supports the findings of Hancock et al. [10]—that the product performance-based attributes such as competence play the leading role in trust formation.

The next phase of the study will incorporate a smart faucet, which adjusts water flow rate and water temperature for users. This new focal product enables physical interactions and behavior observations. As mentioned as a limitation, because the priming process is intended to be subconscious, we didn't validate if the designated emotions were effectively evoked prior to the interaction. It would be ideal if the effectiveness of the primes can be validated or assessed. A different type of priming stimuli, for example, a video clip, can be potentially implemented to magnify the priming effect. In doing so, we expect to paint a more complete picture of the interaction and trust formation between users and products.

This study demonstrates that trust measuring is sensitive to the type of interaction and environment, and according to the work by Hancock et al. [10], it depends on the product type as well. Even though some analogy may exist between the trust formation of the Amazon Echo and that of autonomous vehicles, more research effort is needed when scaling the approach to other products with greater risk. Yet this study does connect the studies in behavioral psychology and trust in automation and provide guidelines to further study with a variety of autonomous products.

This study manipulated user trust to investigate social interactions with autonomous products and trust formation. Overall, the study demonstrates the complex nature of product-trust in a social environment and shows some promise in embedding emotional cues in the product, while the product performance expected by the users is achieved. This learning will potentially facilitate designers to create intentional design with a focus on the affective process in trust formation and to build a healthy level of trust with autonomous products.

ACKNOWLEDGMENTS

This work was funded by the Hasso Plattner Design Thinking Research Program. We thank Hasso Plattner Insitut for supporting this project. We thank Adrienne Kelly Lim and Samantha Beaulieu for their assistance in data collection.

REFERENCES

[1] Hoff, K. A., and Bashir, M., 2015, "Trust in Automation: Integrating Empirical Evidence on Factors that Influence Trust," Hum. Factors, **57**(3), pp. 407-434.

[2] Parasuraman, R., and Riley, V., 1997, "Humans and Automation: Use, Misuse, Disuse, Abuse," Hum. Factors, **39**(2), pp. 230-253.

[3] Borum, R., 2010, "The Science of Interpersonal Trust."

[4] Nass, C., and Moon, Y., 2000, "Machines and Mindlessness: Social Responses to Computers," J. Soc. Issues," **56**(1), pp. 81-103.

[5] Damen, N. B., and Toh, C. A., 2018, "Implicit and Explicit Trust Behavior: Does Stereotype Congruence Affect User Trust in a Home Automation Device?" Proceedings of the ASME 2018 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Quebec City, Canada, Aug. 26-29, pp. V007T06A033-V007T06A033.

[6] Lee, J. D., and See, K. A., 2004, "Trust in Automation: Designing for Appropriate Reliance," Hum. Factors, **46**(1), pp. 50-80.

[7] Friedman, R. S., Fishbach, A., Forster, J., and Werth, L., 2003, "Attentional Priming Effects On Creativity," Creat. Res. J., **15**(2–3), pp. 277-286.

[8] Stone, B., and Soper, S., 2014, "Amazon Unveils a Listening, Talking, Music-Playing Speaker for Your Home". Bloomberg Businessweek. Bloomberg L.P. Accessed Feb. 20, 2019, https://www.bloomberg.com/news/articles/2014-11-06/amazon-echo-is-a-listening-talking-music-playing-speaker-for-your-home

[9] Mayer, R. C., Davis, J. H., and Schoorman, F. D., 1995, "An Integrative Model of Organizational Trust," Acad. Manag. Rev., **20**(3), pp. 709-734.

[10] Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y., Visser, E. J., and Parasuraman, R., 2014, "A Meta-Analysis of Factors Influencing the Development of Trust in Human-Robot Interaction," Hum. Factors, **53**, pp. 517-527.

[11] Nass, C., Steuer, J., and Tauber, E. R., 1994, "Computers Are Social Actors."

[12] Zlotowski, J., Proudfoot, D., Yogeeswaran, K., and Bartneck, C., 2015, “Anthropomorphism: Opportunities and Challenges in Human–Robot Interaction,” Int. J. Soc. Robot., **7**, pp. 347-360.

[13] Epley, N., Waytz, A., and Cacioppo, J. T., 2007, “On Seeing Human: A Three-Factor Theory of Anthropomorphism,” Psychol. Rev., **114**, pp. 484-886.

[14] Lopatovska, I., and Williams, H., 2018, “Personification of the Amazon Alexa: BFF or a Mindless Companion,” Proceedings of the 2018 Conference on Human Information Interaction & Retrieval, New Brunswick, NJ, USA, March 11-15, pp. 265-268.

[15] Culley, K. E., and Madhavan, P., 2013, “A Note of Caution Regarding Anthropomorphism in HCI Agents,” Comput. Hum. Behav., **29**(3), pp. 577-579.

[16] She, J., and MacDonald, E. F., 2018, “Exploring the Effects of a Product’s Sustainability Triggers on Pro-environmental Decision-making,” ASME J. Mech. Des., **140**(1), p. 011102.

[17] She, J., and MacDonald, E. (2014). Priming designers to communicate sustainability. ASME J. Mech. Des., **136**(1), p. 011001.

[18] Liao, T., and MacDonald, E. F., 2018, “Effects of Collage Priming on Sustainable Design Idea Creation and Assessment,” Proceedings of the ASME 2018 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Quebec City, Canada, Aug. 26-29, pp. V007T06A056-V007T06A056.

[19] Bruner, J. S., 1957, “On Perceptual Readiness,” Psychol. Rev., **64**(2), p. 123.

[20] Bargh, J. A., Chen, M., and Burrows, L., 1996, “Automaticity Of Social Behavior: Direct Effects Of Trait Construct And Stereotype Activation On Action,” J. Pers. Soc. Psychol., **71**(2), pp. 230-244.

[21] Lewis, S., Dontcheva, M., and Gerber, E., 2011, “Affective Computational Priming and Creativity,” Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vancouver, Canada, May 7-12, pp. 735-744.

[22] Al-Ubaydli, O., Houser, D., Nye, J., Paganelli, M. P., and Pan, X. S., 2013, “The Causal Effect of Market Priming on Trust: An Experimental Investigation Using Randomized Control,” PloS One, **8**(3), p. e55968.

[23] Callan, M. J., Kay, A. C., Olson, J. M., Brar, N., and Whitefield, N., 2010, “The Effects of Priming Legal Concepts on Perceived Trust and Competitiveness, Self-Interested Attitudes, and Competitive Behavior,” J. Exp. Soc. Psychol., **46**(2), pp. 325-335.

[24] Brownlow, S., 1992, “Seeing is Believing: Facial Appearance, Credibility, and Attitude Change,” J. Nonverbal Behav., **16**(2), pp. 101-115.

[25] Hooker, C. I., Tully, L. M., Verosky, S. C., Fisher, M., Holland, C., and Vinogradov, S., 2011, “Can I Trust you? Negative Affective Priming Influences Social Judgments in Schizophrenia,” J. Abnorm. Psychol., **120**(1), p. 98.

[26] Rice, S., Clayton, K. D., Wells, A., and Keller, D., 2008, “Manipulating Trust Behaviors in a Combat Identification Task.”

[27] Fazio, R. H., and Olson, M. A., 2003, “Implicit Measures in Social Cognition Research: Their Meaning and Use,” Annual Review of Psychology, **54**(1), pp. 297-327.

[28] Yeh, M., and Wickens, C. D., 2000, “Attention and Trust Biases in the Design of Augmented Reality Displays,” Illinois University at Urbana-Champaign Savoy Aviation Research Lab, No. ARL-00-3/FED-LAB-00-1.

[29] Koo, J., Kwac, J., Ju, W., Steinert, M., Leifer, L., and Nass, C., 2015, “Why Did My Car Just Do That? Explaining Semi-Autonomous Driving Actions to Improve Driver Understanding, Trust, And Performance,” Int. J Interact. Des. and Manuf. (IJIDeM), **9**(4), pp. 269-275.

[30] Muir, B. M., and Moray, N., 1996, “Trust in Automation. Part II. Experimental Studies of Trust and Human Intervention in A Process Control Simulation,” Ergonomics, **39**(3), pp. 429-460.

[31] Jian, J. Y., Bisantz, A. M., and Drury, C. G.， 2000, “Foundations for An Empirically Determined Scale of Trust in Automated Systems,” Int. J. Cogn. Ergon., **4**(1), pp. 53-71.

[32] Verberne, F. M., Ham, J., and Midden, C. J., 2012, “Trust in Smart Systems: Sharing Driving Goals and Giving Information to Increase Trustworthiness and Acceptability of Smart Systems in Cars,” Hum. Factors, **54**(5), pp. 799-810

[33] Steinke, F., Ingenhoff, A., and Fritsch, T., 2014, “Personal Remote Assistance in Ambient Assisted Living—Experimental Research of Elderly People’s Trust and Their Intention to Use,” Int. J. Hum. -Comput. Interact., **30**(7), pp. 560-574.

[34] Kurdi, B., Lozano, S., and Banaji, M. R., 2017, “Introducing the Open Affective Standardized Image Set (OASIS),” Behav. Res. Methods., **49**(2), pp. 457-470.

[35] Lang, P. J., Bradley, M. M., and Cuthbert, B. N., 1997, “International Affective Picture System (IAPS): Technical Manual and Affective Ratings,” NIMH Center for the Study of Emotion and Attention, **1**, pp. 39-58.

[36] Bradley, M. M., and Lang, P. J., 2015, “Memory, emotion, and pupil diameter: Repetition of natural scenes,” Psychophysiology, **52**(9), pp. 1186-1193.

[37] Bradley, M. M., Costa, V. D., Ferrari, V., Codispoti, M., Fitzsimmons, J. R., and Lang, P. J., 2015, “Imaging Distributed and Massed Repetitions of Natural Scenes: Spontaneous Retrieval and Maintenance,” Hum. Brain Mapp., **36**(4), pp. 1381-1392.

[38] MacDonald, E. F., and She, J., 2015, “Seven Cognitive Concepts for Successful Eco-design”, J. Clean. Prod., **92**, pp. 23-36.

[39] Merritt, S. M., Heimbaugh, H., LaChapell, J., and Lee, D., 2013, “I Trust It, but I Don’t Know Why. Effects of Implicit Attitudes Toward Automation on Trust in an Automated System,” Hum. Factors, **55**, pp. 520-534.

