AN INTERACTION BETWEEN ANTHROPOMORPHISM AND PERSONALITY ON TRUST IN AUTOMATED SYSTEMS

A thesis submitted in partial fulfillment of the
requirements for the degree of
Master of Science

By

ABRAHAM HASKINS

M.A., Capella University, 2015
B.S., University of Texas at Dallas, 2010
B.A., University of Texas at Dallas, 2010

2021

Wright State University

Wright State University

Graduate School

August 23, 2018

I HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER MY SUPERVISION BY Abraham Haskins ENTITLED An Interaction Between Anthropomorphism and Personality on Trust in Automated Systems BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF Master of Science.

Gary Burns, Ph.D.
Thesis Director

Scott Watamaniuk, Ph.D.
Graduate Program Director

Committee for Final Examination

Ion Juvina, Ph.D.

Nathan Bowling, Ph.D.

Gary Burns, Ph.D.

Barry Milligan, Ph.D.
Vice Provost for Academic Affairs
Dean of the Graduate School

ABSTRACT

Haskins, Abraham. M.S. Department of Psychology, Wright State University, 2021. An Interaction Between Anthropomorphism and Personality on Trust in Automated Systems.

Automated assistance is increasingly being implemented in domains ranging from healthcare to transportation. The reason for the tendency for certain users to trust or mistrust automated assistance has been studied to mixed effect. I examined the effect of anthropomorphism as an independent factor on user trust. In addition, I examined the potential for anthropomorphism to act as a moderator between the personality traits of a user and the trust a user demonstrates in the automated aid. Though the participants in the anthropomorphic condition did view the assistant as more human-like, the level of anthropomorphism had no effect on user behavior. The traits that have been previously found to have an effect on interaction with an automated assistant had their impacts reversed. Users high in extraversion and trait trust were less likely to display trusting behaviors when dealing with an anthropomorphized automated assistant. This expands trait activation theory to the domain of automated interaction. It also allows for a more nuanced understanding of user-automation interaction that impacts selection for any position that interacts with an automated assistant.

Keywords: automation, anthropomorphism, trust, experimental study, trait activation

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

An Interaction Between Anthropomorphism and Personality on Trust in Automated Systems

Researchers and practitioners have been scrambling to increase our understanding of how human behavior changes in response to rapidly increased levels of automation in the business world. Early automation research speculated that trust might influence the relationship between people and automation just as it does between separate people (Sheridan & Hennessy, 1984). More recent research has supported this concept and continued to draw a connection between these two sets of interactions (Lesandowsky, Mundy, & Tan, 2000). Research has shown that many personality factors and automation features correlate with initial trust levels that a user feels towards an automated system (Hoff & Bashir, 2015). For example, extraversion correlates positively with propensity to trust in automated advice, and neuroticism correlates negatively with trusting behaviors in general (Merritt & Ilgen, 2008; Szalma and Taylor, 2011). These results show a parallel with the influence of personality factors on interpersonal trust, albeit with weaker correlations seen within the domain of human-automation trust (Hiraishi, Yamagata, Shikshima, & Ando, 2008). Additionally, research has suggested that anthropomorphizing an automated assistant can change the effect that many traits such as age or gender have on trust levels (Bass, Baumgart, & Shepley, 2013, Pak, Fink, Price, Bass, & Sturre, 2012). Within trait activation theory, this suggests that anthropomorphism may act as a trait-relevant activation cue for at least some relevant interpersonal traits (Tett, Simonet, Walser, & Brown, 2013). However, research has not addressed how personality factors we know to correlate with trust are influenced by the level of anthropomorphism in automated assistance. A user with a high score on an interpersonal scale of cooperation might cooperate more readily with other humans, but does that generalize to cooperation with a human-like machine? Thus, the purpose of my study is to determine how the traits that correlate with trust in automation interact with the addition of an anthropomorphized automated assistant.

Background

The increased use of automation in the business world has led to an accompanying increase in research attempting to predict that automation's effect (Jones, Sundaram & Chin, 2002; Parthasarthy & Sethi, 1992; Venkatraman, 1994). This research has largely focused on practical concerns such as the understanding of human behavior when using an automated vehicle (Carlson, Desai, Drury, Kwak, & Yanco, 2014; Koo et. al, 2015). However, as the use of automation has increased researchers have sought a more general model that predicts the behavior of a user interacting with an automated system. Researchers have proposed several models, many of which have been updated as the research uncovers new information (Heikoop, Winter, Aren, & Stanton, 2016; Lee & See, 2004; Parasuraman, Sheridan, & Wickens, 2000). Adding to that research, my study tests the interaction of two main theoretical determinants of trust in a nonhuman agent: personality factors and the level of anthropomorphism in the automated assistant.

Trust. Interpersonal trust research has a large body of research supporting it, and the most commonly used definition comes from Mayer, Davis, and Shoorman (1995): "the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that party" (p.712). Lee and See (2004) adapted this definition for use with automation as "the attitude that an agent will help achieve an individual’s goals in a situation characterized by uncertainty and vulnerability” (p.51).

The use of this definition is inconsistent throughout the literature, as trust is operationalized in a variety of ways. For example, both Mathur and Reichling (2009) and Beggiato and Krems (2013) used Lee & See’s definition, but both operationalized its measure differently. Mathur & Reichling (2009) operationalized trust through behavior in a game by measuring how much money was given during certain steps. Beggiato and Krems (2013) instead measured changes to state trust in an automated driving simulation through the use of a trust scale. Their usage is not the only one in the research that has led to potential confusion. It is pervasive throughout the body of literature. To address this issue and standardize the definition across the research, some researchers have made specific scales to standardize the definition (Jian, Bisantz, & Drury, 2000).

Operational Trust Measures. The most common operationalization of trust comes from Lee and See (2004), who define automation trust in terms of reliance and compliance. Reliance refers to the behavior of an operator when no alert is given, which signals that all is well. This is further broken down into appropriate reliance and inappropriate reliance. Appropriate reliance describes a lack of action on the part of an operator when an automated assistant correctly rejects the presence of a signal. Inappropriate reliance describes a lack of operator action during a “miss,” wherein the automated assistant fails to detect a signal. Compliance refers to the behavior of an operator when an alert has been issued. Compliance is similarly broken down into appropriate compliance and inappropriate compliance: appropriate compliance referring to action taken by the operator as a result of an alert, and inappropriate compliance referring to action taken by the operator as a result of a "false alarm."

In some studies of automation, including the current study, the constructs of inappropriate and appropriate reliance and compliance are not useful. These constructs, while widely used, fail to consider situations in which the user disagrees with the advice given by the automated assistant. Parasuruman and Riley (1997) addressed this problem by differentiating between use, misuse, disuse, and abuse. Use refers to all situations in which a user asked for advice, received correct advice, and followed the advice given. Misuse refers to situations in which a user asked for advice, received incorrect advice, and followed the advice given. Disuse refers to situations in which a user asked for advice, received incorrect advice, and disagreed with the advice given. Abuse refers to situations in which a user asked for advice, received correct advice, and disagreed with the advice given. These definitions are shown in a table in Figure 1 as applied to the current study, in addition to the mapping of inappropriate and appropriate reliance and compliance for the sake of clarity.

Stages of Trust. Researchers have categorized trust in technology into three groups: *dispositional trust*, the user's predisposition towards being trusting or not; *learned trust*, the user's general tendency to trust the automation as a result of experience; and *situational trust*, the manner in which the user's level of trust changes as a result of the current situational cues in the interaction with the automation (March & Dibben, 2003). In the current study I will examine the effects of personality on dispositional and situational trust, though several of the exploratory questions examined focus on potential interactions in the area of learned trust.

Automation Trust Models

The earliest models describing trust in automation attempted to link trust in automation to trust models found in interpersonal domains (Muir, 1994). Research has upheld this creation of a specific trust model to guide research in the domain of automation trust, because trust in domains within technology appears to have a different set of properties from those found within other domains (Gefen, Karahanna, & Straub, 2003). Corritore, Kracher, and Wiedenbeck (2003), for example, found evidence supporting the idea that trust in online interactions has different properties than trust in other domains. They found that a user's perception of credibility derived from the user interface influences trust levels. While there is the potential analog of physical appearance in interpersonal trust, the predictor of computerized user interface design quality could not exist prior to the use of computers.

Trust models developed further in other domains, and trust models in automation developed alongside them. Mayer et al.'s (1995) trust model is still used in interpersonal research today, and Lee and See used that model in 2004 to develop an updated model for trust in automation. Lee and See's (2004) model is used most often in modern research. March and Dibben's (2003) division into the categories of dispositional, learned, and situational trust often similarly informs the usage of trust in the literature. However, Lee and See's model is often too general for use in subcategories of automated interaction. For example, automated driving has a number of models that describe the interactions that take place within that specific domain (Heikoop, de Winter, Arem, & Stanton, 2016; Stanton & Young, 2000).

Lee and See's (2004) model described several factors that might influence trust in automation and researchers have found evidence for many of its facets. Timeliness of advice, expected levels of accuracy, the presence of a certainty approximation, usage over time, the phrasing of the message, general understanding of how the automated system works, and many other factors have been empirically supported in Lee and See's model (Abe & Richardson, 2005; Beggiato & Krems, 2013; Carlson, Desai, Drury, Kwak, & Yanco, 2014; Helldin, Falkman, Riveiro, & Davidsson, 2013, Kazi, Stanton, Walker, & Young, 2007; Koo, Kwac, Ju, Steinert, Leifer, & Nass, 2015; Piccinini, Rodrigues, Leitao, & Simoes, 2015). In this study I will examine the subset of trust attributes that lead to trust evolution and examine an interaction with the interface features presented in Lee and See's (2004) model.

Dispositional and Situational Trust. Lee and See (2004) do not address the categories of dispositional, situational, and learned trust. These categorizations are used throughout the literature, but they are part of an entirely different model explored by Marsh and Dibbens (2003). The Marsh and Dibbens model is distinct in that it generalizes more readily to other domains (Colquitt, Scott, & LePine, 2007). Dispositional trust is a trait that refers to a user's propensity to trust. Situational trust is a state affected by the properties of the automated system itself. Learned trust is a state affected by interactions with the automated system over time. In this case, I am measuring a potential interaction between an aspect of dispositional trust and situational trust: personality characteristics and the level of anthropomorphism in an automated assistant.

Personality Factors. Research has found that personality traits such as extraversion can correlate with an increase in the level of interpersonal trust displayed (Mooradian, Renzl, & Matzler, 2006). Research has further suggested that whereas there are differences between human-human and human-automation interaction, there exist similar correlations between dispositional trust and some personality features within the domain of human-automation interaction (Madhaven & Wiegmann, 2007). As discussed previously, the exact manner in which these personality factors correlate with dispositional trust is domain-specific and so must be examined independently within the domain of human-automation interaction.

Whereas some research aims at examining the potential influence of personality factors on trust behaviors, it is comparatively sparse when measured against the body of research measuring the factors determined by the properties of the automated aid - such as the degree to which a designer anthropomorphized the automation. However, researchers have uncovered some effects. Merritt and Ilgen (2008) found that users high in extraversion showed a greater propensity to trust automated assistance. Szalma and Taylor (2011) found that users low in neuroticism also showed greater levels of trust in automated assistance. Outside of that, very little research has been aimed in this direction. This is in part due to research by Hancock, Billings, Schaefer, Chen, and De Visser (2011), who found that the impact of human factors on trust in automation is negligible compared to that of the features of the automated assistant. However, they are clear that “the lack of findings may be attributable to insufficient empirical data” in the area of human factors research (Hancock et al., 2011, p. 525). Furthermore, their research showed a high level of variability in the degree to which human attributes influenced propensity to trust the advice of an automated system.

Anthropomorphism. The level of anthropomorphism in an automated system highly influences human-computer interaction. Research has shown this effect to exist across all tested domains of automation (Epley, Waytz, & Cacioppo, 2007; Waytz, Cacioppo & Epley, 2010; Waytz, Heafner, & Epley; 2014). Waytz et al. (2014) suggest that as technology gains human-like mental capabilities, users should increasingly trust it to perform its intended function. Whether this is an inherent property of anthropomorphism has not been researched. For example, anthropomorphism may be a signal to the operator that a particular system is well-designed and thus can be trusted.

Furthermore, there is an interaction effect between many of the traits that correlate with dispositional trust and the level of anthropomorphism seen in an instance of automated assistance. Pak, Fink, Price, Bass, and Sturre (2012) found that as the level of anthropomorphism increased, younger participants were more likely to have increased levels of situational trust than were older participants. The effect gender has on automation trust is also moderated by the level of anthropomorphism displayed in an automated system (Nomura, Kanda, Suzuki, & Kato, 2008; Tung, 2011). This study seeks to extend this research into the effect that personality traits have on automation trust. This would create a parallel in automation trust to the effect of personality on behavior in interpersonal interactions seen in frameworks such as trait activation theory.

When conducting a study measuring differing levels of anthropomorphism, Gong (2008) noted that a method of operationalizing anthropomorphism is a key issue that has not received adequate focus across the literature. To solve this problem, they had participants rate pictures as either more or less human-like in appearance. Their results showed that a human face is rated as significantly more anthropomorphic than a face depicting a stereotypical robot. Researchers have also found that the use of either a synthesized voice or a human voice and the presence of a name for the automated assistant lead to similar assessments of anthropomorphism by participants (Eyssel et al., 2012; Waytz, Heafner, & Epley, 2014; Eyssel & Kuchenbrandt, 2012). This study will utilize this research to provide a name, a human face, and a non-synthesized voice to anthropomorphize an automated assistant.

Trait Activation Theory. This idea that personality traits may be moderated by aspects within a situation is not a new one. Trait activation theory notes that the traits that influence behavior only do so in situations relevant to their activation (Kenrick & Funder, 1988). This assertion, and trait activation theory as a whole, may be relevant to interactions with automated systems (Tett, Simonet, Walser, & Brown, 2013). Within trait activation theory a trait is dormant until it is relevant to the situation at hand, such as the trait of anxiety becoming active when a threatening situation is present. These activating environmental components are called trait-relevant activation cues. Many of the traits involved in interpersonal interaction, such as cooperativeness, are keyed with the trait-relevant activation cue of interaction with another person (Hochwarter, Witt, Treadway, & Ferris, 2006). If a person is anthropomorphizing an automated assistant, this effect could be shown as an increased influence of potentially relevant traits on their behavior. The presence of anthropomorphism could be then treated as a trait-relevant activation cue, encouraging the activation of traits that deal with interpersonal interaction. For example, a trusting individual might demonstrate more trusting behaviors when dealing with a human-like machine than they would when dealing with an assistant that lacks a face or name. If this is true, then it would explain much of the variance seen in the relevance of personality traits in human-automation interaction.

Situational Strength. The situational strength model is another lens through which this moderation effect may be explained (Mischel & Shoda, 1995). Under the situational strength framework, one may inhibit the expression of personality characteristics as a result of being within a strong situation. Within a weak situation, the opposite occurs. Situational strength theory categorizes situations on the basis of four facets: constraints, consequences, clarity, and consistency (Meyer, Dalal, & Hermida, 2010). The presence of an anthropomorphized automated assistant may influence the facets of constraint and consequences. If the user treats the anthropomorphized automated assistant as human, then the constraints that are normally experienced by a person in a strong situation may mirror this interaction with an anthropomorphized decision support aid. Similarly, it is not impossible that at higher levels of anthropomorphism, the consequences of a user's actions upon the automated assistant may be seen as relevant to the user. This would qualify as an application of ethopoeia, as the user would be applying social rules to the interactions with the automated assistant.

Current Study

Research has revealed that certain personality factors, such as extraversion and neuroticism, correlate with dispositional trust levels and automation use (Merrrit & Ilgen, 2008; Szalma & Taylor, 2011). Many of the big five are involved in trust interactions, most notably extraversion and negative neuroticism on the part of the trustor (Evans & Revelle, 2008). While I focus mainly on the big five personality factors, I also adopted a content analysis approach to identify several additional personality traits that might overlap with the decision to use and/or trust an automated aid. Additionally, research has shown that personality factors, despite lacking in meta-analytic data, have a high degree of variability in the influence that they have on trust in an automated system (Hancock et al., 2011). This high level of variability suggests that there are moderator variables that determine the extent that personality is related to trust in automation. One such potential moderator is the level of anthropomorphism of the automated aid.

Anthropomorphism not only has a large effect on the level of trust displayed by a user in human-automation interaction, but is also known to interact with other traits that correlate with dispositional trust such as gender and age (Pak et al., 2012; Nomura et al., 2008; Tung, 2011). However, no one has examined whether anthropomorphism might be a trait activation cue and moderate the influence of personality traits on trust in an automated system. This is important as automation is increasing in commonality in the workplace and across daily life (Jones et al., 2002; Parthasarthy & Sethi, 1992; Vekatraman 1994). Furthermore, Fussell, Kiesler, Selock, and Yew (2008) assert that users inherently anthropomorphize the systems they use over time. This potential interaction between personality traits and anthropomorphism is an important detail to be understood, and as automation permeates further into the business world researchers can only expect its importance to increase.

Hypothesis 1: Increased anthropomorphism of the automated aid will result in a) increased use of the automated aid and b) increased reliance on and compliance with the automated aid (whether inappropriate or appropriate).

Hypothesis 2: Anthropomorphism of the automated aid will moderate the relationships of the selected personality traits with a) use of the automated aid and b) trust in the automated aid such that personality will be more strongly related to appropriate reliance and compliance in the increased anthropomorphism condition.

Method

Participants

Participants were recruited from Wright State University via the SONA system. My initial sample size was 197. 67 participants were removed due to providing incorrect identification numbers and an accompanying failure to match all three parts of the experiment. 4 participants were removed due to a failure to successfully complete the tutorial portion of the experiment. These participants never asked for help from the automated assistant, whether when told to by the experimenter or during the experiment itself. A final sample size of 126 remained. Due to errors in forms filled out by participants, another group of participants had to be excluded from any analysis involving Bartneck, Kulic, and Croft's (2008) Godspeed measure. All analyses made using that measure used a sample of 96. The average age of the participants was 19.7 with a range of 18-37, and it will be made up of 60% women. The sample consisted of 13% African Americans, 2% Asian Americans, 80% Caucasians, 3% Hispanic, and 2% from other ethnicities.

Procedure

First Stage. This study involved three stages of participation for participants. In the first stage, participants were asked to fill out a survey at least 24 hours prior to arriving in the lab. In this survey were tests for 16 personality metrics, demographic questions, and a cognitive test. The 24-hour delay was imposed to prevent fatigue effects from interfering with the responses to the other stages of the study.

Second Stage. In the second stage, participants arrived in our lab to carry out our X-Ray Screening Task. I used an adapted version of the X-Ray Screening Task developed by Merritt and Ilgen (2008). The component images used in the task were obtained from Merrit and Ilgen directly, though they were assembled into a different set of composite images. Both the pilot data and the current study showed similar rates of accuracy with participants as were reported in Merritt and Ilgen's study. Participants' unassisted accuracy was an average of 68% across all images.

In this task, participants were given a set of images such as those one might see at an airport luggage screening device. The participant's goal was to examine each slide consisting of an amalgam of x-ray images and determine if somewhere in this image was a weapon – specifically, a knife or a gun. The task is analogous to the job of a TSA agent screening bags for weapons at an airport. The base rate of targets (weapons) was 50%.

Participants were given unlimited time to scan each image to determine whether a weapon was present. After examining the images shown, participants were instructed to indicate via key response whether a package contained a weapon. A total of 150 images were presented in three blocks of 50. After each block participants were given a chance to rest, informed of their progress for the previous block, and told to continue when ready. This is analogous to a review and subsequent break given by a supervisor to a TSA agent. An image of the stimulus presented to the participants can be found in Figure 2.

The participants were told that there was an automated assistant that could scan for weapons, but that it required oversight and was not always accurate. In actuality, the automated assistant had a flat 80% accuracy rate for the identification of weapons for any image shown. In order to access the assistance of the automated aid, participants were taught that they could press a key to request assistance. A 2-second progress bar was displayed, and a recorded voice played alongside a text response stating that “this package seems to contain a weapon” or “this package is safe.” Each participant was randomly assigned to one of two groups: one with a non-anthropomorphic assistant and one with an anthropomorphic assistant.

Third Stage. The third stage of the study was conducted in the lab immediately following the x-ray screening task. Participants were led to a separate area of the lab and given a written questionnaire with 3 parts. Participants were first asked to write a single sentence to describe their automated assistant. The subject of this sentence was categorized and recorded as either “it,” “he/they/him,” “the assistant,” “she/her,” or “other.” Participants were then asked to estimate both their effectiveness on the task without aid and the effectiveness of the automated assistant. A percentage rating was requested. Participants were then given an adapted version of the Godspeed questionnaire developed by Bartnec, Kulic, and Croft (2009). This questionnaire is designed to assess the perceived level of anthropomorphism of an automated system. It was adapted for both language and purpose by replacing inapplicable prompts such as “moves rigidly/moves elegantly” with “speaks rigidly/speaks elegantly.”

Manipulation

Participants were randomly assigned to one of two groups. One group had an anthropomorphized automated assistant, and one had a non-anthropomorphized automated assistant. The anthropomorphic and the non-anthropomorphic assistant were functionally identical. Each had identical methods of access, response phrasing, and accuracy ratings. The assistants differed in three ways: name, voice, and picture.

Both groups of participants shared the same room. Viewing another computer was not possible due to large dividers, and the auditory stimuli were delivered via headphones. During the initial training portion of the study, the experimenters were trained to respond to questions using neutral language, referring to the assistant as “the assistant” rather than through the use of a pronoun or a reference to the name of the assistant.

Non-Anthropomorphic Assistant. The non-anthropomorphic assistant was given the initialization AWD, which stood for Automated Weapons Detector. Its verbal response to a request for assistance consisted of a synthesized voice. A small picture of a computer was used as a replacement for the picture of a person used in the anthropomorphic condition.

Anthropomorphic Assistant. The anthropomorphic assistant was given the acronym “AWDi,” which stood for Automated Weapons Detector, Interactive. It was given this name due to the ability for this name to be pronounced as a complete name rather than an initialization. In addition, the verbal response to a request for assistance used a verbal response recorded by a volunteer who had no interaction with any of the participants. Last, the anthropomorphic assistant had a small picture of a man in the area of the screen designated for the automated assistant.

Participant Motivation

A participant's results included a metric for accuracy and the participants were made aware of this at the start of the study. The participants were also told that the most accurate scorers would receive a $50 gift card. This compensation was intended as a motivating force that would prevent low-effort responding.

Measures

Personality Measures. Many of the big five traits, most notably extraversion and neuroticism, have been found to correlate with interpersonal trust (Mooradin, Renzl, & Matzler, 2006). The big five factors were measured with 10-item scales drawn from the International Personality Item Pool (Goldberg et al., 2006). The Cronbach's alpha of the scales used ranged from .79 to .87. While extraversion and negative neuroticism are both related to interpersonal trust, many facets and blends of the big five are also correlated with interpersonal trust (Mooradin, Renzl, & Matzler, 2006). Assertiveness, competence, and trait trust scales were included for this reason. These measures were drawn from the IPIP AB5C measures and have Cronbach's alpha estimates ranging from .74 to .82.

Due to the nature of the current study, a number of other traits might be relevant in this particular interaction. I adopted a content analysis approach to identify several additional personality traits that might relate to the decision to use and/or trust an anthropomorphized automated aid. For example, the scale of dominance from the IPIP includes items such as "I impose my will on others" (Goldberg et al., 2006). This is a trait that might potentially be moderated by the presence of a more human-like automated assistant. This trait may affect a participant's behavior with a human assistant, but not necessarily with an automated assistant. An anthropomorphized assistant may act as a moderator on the effect of this trait on trusting behavior.

Measures of 10 potentially relevant traits were included with the big five in the initial questionnaire. Each measure taps into a trait that may have a different effect on trusting behavior depending on whether the interaction is with a human or an automated assistant. Conformity may only apply when the participant is prone to conform with another human. Cooperation may not apply when receiving information and advice from an inhuman tool. The measure of flexibility includes items such as "I am good at taking advice," which implies a human actor giving such advice. Self-sufficiency may not apply when working with an automated aid, as it may be treated as a tool rather than an outside influence from which it is possible to be independent. Perfectionism includes items such as "I demand perfection in others," which may impact how readily advice is taken from an occasionally incorrect human more so than an occasionally incorrect automated assistant. Rigidity similarly has many items that refer to a willingness to change due to the influence of the opinions of other humans. Each of these measures has the potential to have a different effect on trusting behavior depending on whether it is a human or an automated assistant giving advice.

Each of these measures were taken from the IPIP (Goldberg et al., 2006). The traits selected and their respective reliabilities are assertiveness ($\alpha = .75$), competence ($\alpha = .74$), conformity ($\alpha = .79$), cooperation ($\alpha = .73$), dominance ($\alpha = .82$), flexibility ($\alpha = .73$), perfectionism ($\alpha = .76$), rigidity ($\alpha = .77$), self-sufficiency ($\alpha = .59$), and trust ($\alpha = .82$). This resulted in 175 items rated on a 5-point graphical rating scale with responses ranging from “describes me” to “does not describe me.” A high score on any of the individual metrics will indicate an increased level in the participant of the trait measured.

Cognitive Test. As part of the pre-test, a cognitive assessment was given to all participants. The Verbal Comprehension and Numerical Ability assessments of the Employee Aptitude Survey were given to all participants (Ruch, Stang, McKillip, & Dye, 1994). The Employee Aptitude Survey has an overall internal consistency reliability estimate of $\alpha = .88$. The Numerical Ability assessment consists of 75 mathematical questions such as “1.2 x 12” and “6% of 10.” Each question was presented in multiple-choice format with five options. The Verbal Comprehension assessment consists of 25 questions instructing the participants to select the synonyms of words such as “meander” and “spasmodic.” Participants were instructed to choose the correct answer out of four available possibilities. Correct answers from the participants on each of the scores were summed. The participants’ correct answer sums were then normalized across the sample group. A high score indicated an increased cognitive score.

Experimental Measures. The four scales of use, misuse, disuse, and abuse were first calculated from the data for each participant. For use, I summed the number of times that a participant responded with “contains a weapon” after being advised by the automated assistant that a weapon existed while dealing with a stimulus in which a weapon was actually present. I added to that the number of times a participant responded with “does not contain a weapon” after being advised by the automated assistant that a weapon was not present while dealing with a stimulus in which a weapon was not actually present. I then divided this sum by the number of times in which a participant requested help. I repeated this process for the other three operational variables: misuse, disuse, and abuse. See Figure 1 for the appropriate mapping of each onto the current study. I then combined use and misuse, creating a variable called "Operational Trust." This variable shows the percentage of times users agreed with either correct or incorrect advice.

Results

Manipulation Check

A t-test was conducted comparing the results of Bartneck, Kulic, and Croft's (2008) Godspeed measure of anthropomorphism to check the effectiveness of my manipulation. A significant difference was found ($t = 2.60$, $df = 89$, $p = .01$). Participants perceived the anthropomorphized automated assistant as actually being more human-like.

Descriptive Statistics and Correlations

Table 1 reports the means and standard deviations for all experimental variables. Table 2 shows correlations between the operational variables and personality factors. Several interesting correlations can be observed in this table. First, no personality traits were correlated with help requests, but cognitive ability was. Second, no individual difference variables were correlated with operational trust. Among the four behavior indicators, some interesting patterns were discovered. For example, Openness shows a negative correlation with Use rates, Perceived Anthropomorphism shows a positive correlation with Assertiveness, Dominance shows a negative correlation with Misuse rates, and Conformity shows a positive correlation with both Use and Disuse rates. This is inconsistent with what would be expected from previous research that found a positive correlation between operational trust and extraversion (Merritt & Ilgen, 2008).

Effect of Anthropomorphism on Usage

In Hypothesis 1a and 1b, I predicted that the participants in the anthropomorphic condition would display a significantly higher level of both requests for help and operational trust across all three task blocks. To test this, I conducted a set of t-tests. The results of a t-test did not reveal a significant difference between the anthropomorphic and non-anthropomorphic conditions for either the help request metric (t = .49, df = 125, p = .62) or the operational trust metric (t = 1.04, df = 125, p = .30). Similar nonsignificant results were found when examining the results of t-tests solely within the first block of trials. Testing this hypothesis by examining the degree to which a given participant anthropomorphized the automated assistant instead of that participant's experimental condition yielded a different pattern of results. To examine this, I conducted a series of linear regression analyses. Regressing the help request metric on the results of the Godspeed measure of anthropomorphism showed a negative relationship (β = -.86, p = .005, df = 92). This negative relationship indicated that participants who perceived the automation as more anthropomorphic, regardless of the condition, tended to use the automation less. Regressing the operational trust metric onto the Godspeed measure showed no relationship (β = .09, p = .21, df = 92).

Anthropomorphism's Moderating Effects on Personality

In Hypotheses 2a and 2b, I hypothesized that anthropomorphism would moderate the effects of personality's influence on help request frequency and operational trust such that increased anthropomorphism would strengthen the relationship. In order to test these hypotheses a series of moderated regression analyses were conducted (see Table 3). The significant relationships in the first steps of these regressions mirrored those found in the correlation run in the initial data examination.

Largely, there was no moderation. In the 30 analyses, only three interaction effects were significant. Openness, extraversion, and trust interacted with the manipulation to predict operational trust and help request frequency. However, effects were found to be in the opposite direction as was predicted. Rather than increasing anthropomorphism leading to a more positive effect of personality on behavior, an increase in anthropomorphism led to a negative relationship. These relationships are illustrated in Figures 3, 4, and 5. The other variable known to affect automation use, neuroticism, had interaction terms significant only at the $p < .1$ level.

Discussion

This research aimed to provide support for the positive effect of anthropomorphism on trust behaviors within a novel domain and to show that the presence of anthropomorphism moderated the relationship between personality traits and trust behaviors. I did not find evidence for an effect on behavior by anthropomorphism, even though the participants saw the anthropomorphized assistant as more human-like. The moderating effect of anthropomorphism, when present in 10% of analyses, gave the opposite result as was expected. The presence of anthropomorphism led to a decrease in the relationship of extraversion, openness, and interpersonal trust with operational trust.

The results from the check of the first hypothesis run counter to what has been observed in other experiments involving anthropomorphism. For example, Waytz et. al.'s (2014) research involving anthropomorphized driving assistants found a significant positive relationship between operational trust and the level of anthropomorphism. This indicates that there is some critical difference between other studies and this one that nullifies the effect of anthropomorphism. The results from the check of the second hypothesis provide some support for the application of trait activation theory within the domain of automation interaction. Participants that were presented with what they perceived as something human-like responded to it as they would with the presence of another person. Those participants slightly repressed their relevant personality traits.

The results from this experiment are somewhat unique. There are three potential explanations for these results. The first is that there is a domain restriction on the effect of anthropomorphism that has not been previously uncovered. A novel and difficult task that would require training and expertise in the real world, examining x-rays of luggage, was carried out by completely untrained individuals in an unfamiliar environment. Any one of those factors, or a combination thereof, may have led to a unique environment that prevents the expected effects of anthropomorphism from manifesting. The second potential explanation relates to the specifics of the program. For example, the automated assistant in the anthropomorphic condition did not look like an expert in the field of baggage screening. This or another factor may have led to an overall decrease in trust that countered the otherwise positive effects that anthropomorphism has on usage behaviors. The third potential explanation is that, in this case, the perceived competence of the automated system surpassed that of the anthropomorphized assistant. This would support the research showing that users tend to trust automation over humans in most assessments of competence (Tseng & Fogg, 1999)

The results from the check of the second hypothesis are unintuitive. However, they make sense within the context of trait activation theory. The presence of a positive effect on operational trust by extraversion, for example, has been found before. However, with the addition of another “person” to the activity, this expression of personality is repressed so much as to be reversed. This effect is expected within the framework of trait activation theory. Under trait activation theory, personality traits are increasingly repressed as more people are present in a given environment (Tett, Simonet, Walser, & Brown, 2013). That the “person” referred to here is an anthropomorphized automated assistant merely provides additional evidence for the concept of ethopoeia (Nass & Moon, 2000). This trend is repeated for the behavior displayed by openness and neuroticism.

Theoretical Implications

The results from the test on the first hypothesis indicate that the positive effects of anthropomorphism on usage behaviors may be counteracted by at least one factor that has not been previously accounted for. There is some research examining the impact of perceived expertise displayed by the anthropomorphized personality. That research does not claim that the lack of displayed expertise would eliminate the positive effects of anthropomorphism, but that was potentially demonstrated here.

Alternatively, the reason for this lack of effect by anthropomorphism may be a simple domain restriction. Some unknown factors within the study may have reduced the effect of anthropomorphism on behavior. If this is the case, then further research is required to uncover exactly what made this occur.

The findings from my second hypothesis provide evidence in favor of a lack of domain restriction on trait activation theory. This means that trait activation theory may be applied to settings in which the participant is not accompanied by another person at all. Merely the feeling that a person-like entity is nearby may be enough to trigger this change in behavior. The findings also provide an example of a general application of ethopoeia; in this case, to trait activation theory.

Practical Implications

The most relevant application of the results from the first hypothesis is an understanding that merely adding an anthropomorphized automated assistant is not sufficient to increase help requests or operational trust. It is entirely possible for an engineer to add an automated assistant with certain features or within a specific domain that has no effect whatsoever upon usage behaviors. This could lead to more efficient time spent developing anthropomorphized assistants and prevent waste when such an assistant would have no effect.

The results from the second hypothesis are as useful as was expected, but merely need must be applied in reverse. If a truck fleet overseer is needed to regularly question the automated assistant's behavior, then a trusting worker would be more effective with an automated assistant while a suspicious worker would be more effective without. As more jobs open up involving interaction with automated assistance, this helps craft more clear job analyses that would find the right worker for a position that interacts with these assistants.

Future Research

Future research should focus first on identifying what caused the lack of an effect from the presence of anthropomorphism in the automated assistant. Knowing what properties in either the situation or the assistant itself lead to increased changes in behavior would aid both theoretical research and the engineering of new anthropomorphic assistants. An examination of the power of a lack of perceived expertise on the part of the persona displayed by the anthropomorphized automated assistant would be a good starting place.

In addition, this is the first major addition to the list of social rules that may be applied to human interactions with automated systems since the introduction of the concept of ethopoeia in 2000. This means that there are likely many other examples of such social rule applications that have gone unnoticed. Identifying more social norms that apply within this domain would allow for a more complete understanding of what kinds of social norms are generalizable to the domain of human-automation interaction.

Limitations

The main limitation of this study was that of sample size. While an initial size of 150 was required to detect the expected effect sizes, the multi-part nature of the study meant that over a third of all participant data could not be included in the analysis as a result of missing at least one part of the experiment. This limitation was compounded by the nature of the experiment. Out of 150 stimuli presented, the median participant requested help 68 times. Of those 68 requests for help, the median participant disagreed with the automated assistant only 12 times. Many participants never even disagreed at all. In future studies of use, misuse, disuse, and abuse, a much larger stimulus set would be necessary to recognize the specifics of user behavior. In addition to sample size, the participants used in this experiment were largely homogenous in age, ethnicity, and inexperience with the task.

The other limitation to this study is its unclear domain restriction. This study has given an example of a case in which the presence of anthropomorphism did not influence usage behavior. That indicates an unknown domain restriction or program property that negates the expected effect of anthropomorphism. However, this domain restriction is unknown, which makes its application impossible. Knowing that such a domain restriction exists, however, is important information in the application of anthropomorphism to modern automated assistants.

Conclusion

This study set out to provide additional evidence for the benefits of anthropomorphism in automated assistants. Instead, it provided an example of the type of situation in which anthropomorphism did not affect user behavior. The commonality of anthropomorphism in automation is driven in part by our belief in their effectiveness in increasing use rates. If some specific features or domains cause this effectiveness to be eliminated, then it is important that researchers examine more closely what limitations exist. This study also intended to show that the presence of anthropomorphism strengthened the effect of personality characteristics on user behavior. Instead, it showed that the opposite happened. People expressed their personality characteristics less when in the presence of anthropomorphic assistance. While counterintuitive, these results align with what one would expect from personality research. Individuals moderate their behaviors when in the presence of others. Given the success of applying social rules to anthropomorphized systems, it comes as no surprise that this rule would apply as well. Automated aids are becoming nearly ubiquitous in today's world. Many businesses use them, and most people carry one around daily in their pocket in the form of a cell phone. This research shows that not only are some user interactions counter-intuitive, but that even the most basic understanding of anthropomorphic systems use is more complicated than is currently understood.

References

Abe, G., & Richardson, J. (2005). The influence of alarm timing on braking response and driver trust in low speed driving. *Safety Science*, *43*(9), 639-654.

Bass, E. J., Baumgart, L. A., & Shepley, K. K. (2013). The effect of information analysis automation display content on human judgment performance in noisy environments. Journal of cognitive engineering and decision making, 7(1), 49-65.

Beggiato, M., & Krems, J. F. (2013). The evolution of mental model, trust and acceptance of adaptive cruise control in relation to initial information. *Transportation research part F: traffic psychology and behaviour*, *18*, 47-57.

Carlson, M. S., Desai, M., Drury, J. L., Kwak, H., & Yanco, H. A. (2014, March). Identifying factors that influence trust in automated cars and medical diagnosis systems. In *AAAI Symposium on The Intersection of Robust Intelligence and Trust in Autonomous Systems* (pp. 20-27).

Colquitt, J. A., Scott, B. A., & LePine, J. A. (2007). Trust, trustworthiness, and trust propensity: a meta-analytic test of their unique relationships with risk taking and job performance. *Journal of applied psychology*, *92*(4), 909.

Culley, K. E., & Madhavan, P. (2013). A note of caution regarding anthropomorphism in HCI agents. *Computers in Human Behavior*, *29*(3), 577-579.

Corritore, C. L., Kracher, B., & Wiedenbeck, S. (2003). On-line trust: concepts, evolving themes, a model. *International journal of human-computer studies*, *58*(6), 737-758.

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: a three-factor theory of anthropomorphism. *Psychological review*, *114*(4), 864.

Evans, A. M., & Revelle, W. (2008). Survey and behavioral measurements of interpersonal trust. *Journal of Research in Personality*, *42*(6), 1585-1593.

Eyssel, F., & Kuchenbrandt, D. (2012). Social categorization of social robots: Anthropomorphism as a function of robot group membership. *British Journal of Social Psychology*, *51*(4), 724-731.

Eyssel, F., Kuchenbrandt, D., Bobinger, S., de Ruiter, L., & Hegel, F. (2012, March). 'If you sound like me, you must be more human': on the interplay of robot and user features on human-robot acceptance and anthropomorphism. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction* (pp. 125-126). ACM.

Faul, F., E. Erdfelder, A.G. Lang, and A. Buchner (2007), "G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences," *Behavior Research Methods*, 39(2): 175-191.

Fussell, S. R., Kiesler, S., Setlock, L. D., & Yew, V. (2008, March). How people anthropomorphize robots. In *Human-Robot Interaction (HRI), 2008 3rd ACM/IEEE International Conference on* (pp. 145-152). IEEE.

Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and TAM in online shopping: an integrated model. *MIS quarterly*, *27*(1), 51-90.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. C. (2006). The International Personality Item Pool and the future of public-domain personality measures. *Journal of Research in Personality, 40,* 84-96.

Gong, L. (2008). How social is social responses to computers? The function of the degree of anthropomorphism in computer representations. Computers in Human Behavior, 24(4), 1494-1509.

Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. Science, 315(5812), 619-619.

Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y., De Visser, E. J., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors*, *53*(5), 517-527.

Heikoop, D. D., de Winter, J. C., van Arem, B., & Stanton, N. A. (2016). Psychological constructs in driving automation: a consensus model and critical comment on construct proliferation. *Theoretical Issues in Ergonomics Science*, *17*(3), 284-303.

Helldin, T., Falkman, G., Riveiro, M., & Davidsson, S. (2013, October). Presenting system uncertainty in automotive UIs for supporting trust calibration in autonomous driving. In *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 210-217). ACM.

Hiraishi, K., Yamagata, S., Shikishima, C., & Ando, J. (2008). Maintenance of genetic variation in personality through control of mental mechanisms: A test of trust, extraversion, and agreeableness. *Evolution and Human Behavior, 29*(2), 79-85.

Hochwarter, W. A., Witt, L. A., Treadway, D. C., & Ferris, G. R. (2006). The interaction of social skill and organizational support on job performance. Journal of Applied Psychology, 91(2), 482.

Hoff, K. A., & Bashir, M. (2015). Trust in automation integrating empirical evidence on factors that influence trust. Human Factors: The Journal of the Human Factors and Ergonomics Society, 57(3), 407-434.

Jian, J. Y., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an empirically determined scale of trust in automated systems. International Journal of Cognitive Ergonomics, 4(1), 53-71.

Johnson, J. A. (1994). Clarification of factor five with the help of the AB5C model. European Journal of Personality, 8(4), 311-334.

Jones, E., Sundaram, S., & Chin, W. (2002). Factors leading to sales force automation use: A longitudinal analysis. Journal of Personal Selling & Sales Management, 22(3), 145-156.

Kazi, T. A., Stanton, N. A., Walker, G. H., & Young, M. S. (2007). Designer driving: drivers' conceptual models and level of trust in adaptive cruise control. *International journal of vehicle design*, *45*(3), 339-360.

Kenrick, D. T., & Funder, D. C. (1988). Profiting from controversy: Lessons from the person-situation debate. American psychologist, 43(1), 23.

Koo, J., Kwac, J., Ju, W., Steinert, M., Leifer, L., & Nass, C. (2015). Why did my car just do that? Explaining semi-autonomous driving actions to improve driver understanding, trust, and performance. *International Journal on Interactive Design and Manufacturing (IJIDeM)*, *9*(4), 269-275.

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. Human Factors: The Journal of the Human Factors and Ergonomics Society, 46(1), 50-80.

Lewandowsky, S., Mundy, M., & Tan, G. (2000). The dynamics of trust: Comparing humans to automation. *Journal of Experimental Psychology: Applied*, *6*(2), 104.

Madhavan, P., & Wiegmann, D. A. (2007). Similarities and differences between human–human and human–automation trust: an integrative review. *Theoretical Issues in Ergonomics Science*, *8*(4), 277-301.

Mathur, M. B., & Reichling, D. B. (2009, March). An uncanny game of trust: social trustworthiness of robots inferred from subtle anthropomorphic facial cues. In *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction* (pp. 313-314). ACM.

Marsh, S., & Dibben, M. R. (2003). The role of trust in information science and technology. *Annual Review of Information Science and Technology*, *37*(1), 465-498.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. Academy of management review, 20(3), 709-734.

Meyer, R. D., Dalal, R. S., & Hermida, R. (2010). A review and synthesis of situational strength in the organizational sciences. *Journal of Management*, *36*(1), 121-140.

Merritt, S. M., & Ilgen, D. R. (2008). Not all trust is created equal: Dispositional and history-based trust in human-automation interactions. *Human Factors*, *50*(2), 194-210.

Mooradian, T., Renzl, B., & Matzler, K. (2006). Who trusts? Personality, trust and knowledge sharing. *Management learning*, *37*(4), 523-540.

Personality factors can lead to a change in the level of interpersonal trust

Muir, B. M. (1994). Trust in automation: Part I. Theoretical issues in the study of trust and human intervention in automated systems. Ergonomics, 37(11), 1905-1922.

Nass, C., & Lee, K. M. (2001). Does computer-synthesized speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction. Journal of experimental psychology: applied, 7(3), 171.

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of social issues, 56*(1), 81-103.

Nomura, T., Kanda, T., Suzuki, T., & Kato, K. (2008). Prediction of human behavior in human--robot interaction using psychological scales for anxiety and negative attitudes toward robots. IEEE transactions on robotics, 24(2), 442-451.

Pak, R., Fink, N., Price, M., Bass, B., & Sturre, L. (2012). Decision support aids with anthropomorphic characteristics influence trust and performance in younger and older adults. Ergonomics, 55(9), 1059-1072.

Parthasarthy, R., & Sethi, S. P. (1992). The impact of flexible automation on business strategy and organizational structure. Academy of Management review, 17(1), 86-111.

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *39*(2), 230-253.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. IEEE Transactions on systems, man, and cybernetics-Part A: Systems and Humans, 30(3), 286-297.

Piccinini, G. F. B., Rodrigues, C. M., Leitão, M., & Simões, A. (2015). Reaction to a critical situation during driving with adaptive cruise control for users and non-users of the system. *Safety science*, *72*, 116-126.

Ruch, W. W., Stang, S. W., McKillip, R. H., & Dye, M. (1994). Employee Aptitude Survey: Technical manual.. Glendale, CA: Psychological Services.

Schaefer, K. E., Chen, J. Y., Szalma, J. L., & Hancock, P. A. (2016). A Meta-Analysis of Factors Influencing the Development of Trust in Automation Implications for Understanding Autonomy in Future Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 0018720816634228.

Sheridan, T. B., & Hennessy, R. T. (1984). *Research and modeling of supervisory control behavior. Report of a workshop*. National Research Council Washington DC Committee on Human Factors.

Singh, Indramani L., Robert Molloy, and Raja Parasuraman. "Automation-induced" complacency": Development of the complacency-potential rating scale." *The International Journal of Aviation Psychology* 3.2 (1993): 111-122.

Stanton, N. A., & Young, M. S. (2000). A proposed psychological model of driving automation. *Theoretical Issues in Ergonomics Science*, *1*(4), 315-331.

Szalma, J. L., & Taylor, G. S. (2011). Individual differences in response to automation: the five factor model of personality. *Journal of Experimental Psychology: Applied*, *17*(2), 71.

Tett, R. P., Simonet, D. V., Walser, B., & Brown, C. (2013). Trait activation theory. *Handbook of personality at work*, 71-100.

Tseng, S., & Fogg, B. J. (1999). Credibility and computing technology. *Communications of the ACM*, *42*(5), 39-44.

Tung, F. W. (2011, July). Influence of gender and age on the attitudes of children towards humanoid robots. In International Conference on Human-Computer Interaction (pp. 637-646). Springer Berlin Heidelberg.

Venkatraman, N. (1994). IT-enabled business transformation: from automation to business scope redefinition. Sloan management review, 35(2), 73.

Waytz, A., Cacioppo, J., & Epley, N. (2010). Who sees human? The stability and importance of individual differences in anthropomorphism. *Perspectives on Psychological Science, 5*(3), 219-232.

Waytz, A., Heafner, J., & Epley, N. (2014). The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. *Journal of Experimental Social Psychology*, *52,* 113-117.

Table 1

Descriptive statistics

Study Variables	*M*	*SD*
Help Request Frequency	74.17	45.47
%Use	73.53	9.33
%Misuse	16.80	7.00
%Disuse	3.53	4.58
%Abuse	6.14	6.59
%Operational Trust	90.33	10.16
Anthropomorphism Assessment	42.56	15.86
Self-Efficacy Assessment	48.59	21.43
Aid-Efficacy Assessment	62.64	25.19
Cognitive Score	15.35	10.58

Note. N = 126

Table 2 - *Correlation Table for All Study Variables*

Study Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1. Openness																						
2. Conscientiousness	.00																					
3. Extraversion	.31*	.25*																				
4. Agreeableness	.29*	.38*	.40*																			
5. Neuroticism	.05	-.42*	-.37*	-.43*																		
6. Assertiveness	.09	.29*	.59*	.23*	-.31*																	
7. Competence	-.02	.73*	.34*	.33*	-.52*	.48*																
8. Cooperation	-.16*	.19*	-.15	.34*	-.27*	-.16*	.09															
9. Perfectionism	-.04	.52*	.08	.04	-.06	.35*	.58*	-.10														
10. Rigidity	-.37*	-.24*	.04	-.36*	.14	.16*	-.06	-.36*	.13													
11. Conformity	-.08	-.25*	.03	-.16*	.42*	-.09	-.24*	-.25*	-.01	.42*												
12. Dominance	-.16*	.07	.25*	-.33*	-.06	.48*	.22*	-.46*	.31*	.58*	.13											
13. Flexibility	.29*	.22*	.22*	.53*	-.48*	.03	.25*	.46*	-.16*	-.67*	-.39*	-.42*										
14. Self-Sufficiency	.16*	.06	-.01	.02	-.28*	.09	.09	-.14	-.02	-.13	-.46*	-.05	.15									
15. Trust	.26*	.22*	.54*	.59*	-.47*	.41*	.37*	.09	.05	-.11	.08	.04	.43*	.02								
16. Cognitive Ability	.18*	.12*	-.09*	-.03	.04	.02	.07	.00	.07	-.26*	-.14	-.01	.10	.04	.08							
17. Perceived Anthropomorphism	.12	-.04	.07	.07	.04	.22*	.05	-.09	.02	.09	.08	.02	-.06	.03	.07	.12						
18. Help Requests	.01	.01	.03	-.09	.06	-.04	-.03	-.01	.04	-.05	-.01	.04	.07	.03	-.10	.23*	-.15					
19. %Use	-.18*	.09	-.07	-.03	.03	-.02	.08	.05	.16	.11	.19*	.01	-.06	-.11	-.03	-.06	.05	-.17*				
20. %Misuse	.01	.04	.01	.12	.03	-.09	-.06	.11	-.10	-.05	-.05	-.20*	.01	.07	.07	-.05	.11	-.08	-.18*			
21. %Disuse	.14	-.04	.03	.00	-.06	-10	.02	-.08	.00	-.09	-.16*	.09	.04	.03	-.03	.09	-.15	.15	-.71*	-.45*		
22. %Abuse	.14	-.14	.07	-.08	-.04	.06	-.07	-.13	-.11	-.04	-.10	.14	.05	.07	-.01	.07	-.08	.22*	-.73*	-.49*	.79*	
23. Operational Trust	-.13	.06	-.03	.10	.06	-.04	-.03	.10	-.01	.06	.10	-.17	.04	.02	-.02	-.08	.11	-.32*	.74*	.45*	-.92*	-.97*

Note. $N = 126$, $*p < .05$

Table 3

Moderated Regression Analyses Results

Criterion	*Ordered Predictors*	*β*	*Step-2 β*	*Step-3 β*	R^2
Operational Trust	Openness	-.12	-.11	-.11	.02
	A		-.08	-.07	.02
	Openness * A			-.18	.06
Help Requests	Openness	.08	.09	.08	.01
	A		-.05	-.05	.01
	Openness * A			-.07*	.01
Operational Trust	Conscientiousness	.06	.05	-.04	.00
	A		-.09	-.09	.01
	Conscientiousness * A			.05	.01
Help Requests	Conscientiousness	-.02	-.02	-.02	.00
	A		-.05	-.05	.00
	Conscientiousness * A			.04	.00
Operational Trust	Extraversion	-.03	-.02	.01	.00
	A		-.09	-.09	.01
	Extraversion * A			-.29*	.01
Help Requests	Extraversion	.07	.07	.06	.00
	A		.07	-.04	.01
	Extraversion * A			.16	.03
Operational Trust	Agreeableness	.10	.10	.08	.01
	A		-.10	-.10	.02
	Agreeableness * A			-.09	.03
Help Requests	Agreeableness	-.05	-.05	-.02	.00
	A		-.04	0.04	.00
	Agreeableness * A			.13	.02
Operational Trust	Neuroticism	.06	.07	.02	.00
	A		-.10	-.10	.01
	Neuroticism * A			.13	.03
Help Requests	Neuroticism	.02	.03	.08	.00
	A		.03	-.05	.00
	Neuroticism * A			-.15	.02

*Note. N = 126, Each regression run in three steps. All variables normalized. *p < .05*

Table 3 (cont)

Moderated Regression Analyses Results, AB5C Variables

Criterion	*Ordered Predictors*	*β*	*Step-2 β*	*Step-3 β*	R^2
Operational Trust	Assertiveness	-.04	-.03	-.04	.00
	A		-.09	-.09	.01
	Assertiveness * A			-.14	.03
Help Requests	Assertiveness	-.05	-.05	-.04	.00
	A		-.04	-.04	.00
	Assertiveness * A			.09	.01
Operational Trust	Competence	-.03	-.04	-.03	.00
	A		-.10	-.10	.01
	Competence * A			-.04	.01
Help Requests	Competence	-.06	-.07	-.09	.00
	A		-.05	-.05	.01
	Competence * A			.14	.03
Operational Trust	Cooperation	.10	.10	.10	.01
	A		-.10	-.09	.02
	Cooperation * A			.02	.02
Help Requests	Cooperation	.02	.02	.02	.00
	A		-.04	-.04	.00
	Cooperation * A			.08	.01
Operational Trust	Perfectionism	-.02	-.03	-.03	.00
	A		-.10	-.10	.01
	Perfectionism * A			.00	.01
Help Requests	Perfectionism	-.05	-.06	-.06	.00
	A		-.05	-.05	.01
	Perfectionism * A			.01	.01
Operational Trust	Rigidity	.06	.04	.04	.00
	A		-.09	-.09	.01
	Rigidity * A			-.02	.01
Help Requests	Rigidity	-.14	-.16	-.16	.02
	A		-.07	-.07	.03
	Rigidity * A			.11	.04

*Note. N = 126, Each regression run in three steps. All variables normalized. *p < .05*

Table 3 (cont.)

Moderated Regression Analyses Results, Exploratory Variables

Criterion	*Ordered Predictors*	*β*	*Step-2 β*	*Step-3 β*	R^2
	Conformity	.10	.11	.11	.01
Operational Trust	A		-.11	-.11	.02
	Conformity * A			.02	.02
	Conformity	-.03	-.02	-.01	.00
Help Requests	A		-.04	-.04	.00
	Conformity * A			-.05	.00
	Dominance	-.16	-.17	-.17	.03
Operational Trust	A		-.10	-.10	.04
	Dominance * A			-.06	.04
	Dominance	.00	.00	.00	.00
Help Requests	A		-.04	-.04	.00
	Dominance * A			.00	.00
	Flexibility	-.04	-.03	-.03	.00
Operational Trust	A		-.09	-.09	.01
	Flexibility * A			-.10	.02
	Flexibility	.10*	.10*	.18*	.03
Help Requests	A		-.06	-.06	.04
	Flexibility * A			-.00	.04
	Self-Sufficiency	.02	.01	-.02	.00
Operational Trust	A		-.10	-.10	.01
	Self-Sufficiency * A			-.11	.02
	Self-Sufficiency	.05	.04	.01	.00
Help Requests	A		-.04	-.04	.00
	Self-Sufficiency * A			-.11	.02
	Trust	-.02	-.00	-.01	.00
Operational Trust	A		-.10	-.09	.01
	Trust * A			-.22*	.06
	Trust	-.10	-.10	-.11	.01
Help Requests	A		-.03	-.03	.01
	Trust * A			-.06	.02

*Note. N = 126, Each regression run in three steps. All variables normalized. *p < .05*

Figure 1

Operational Variable Mapping

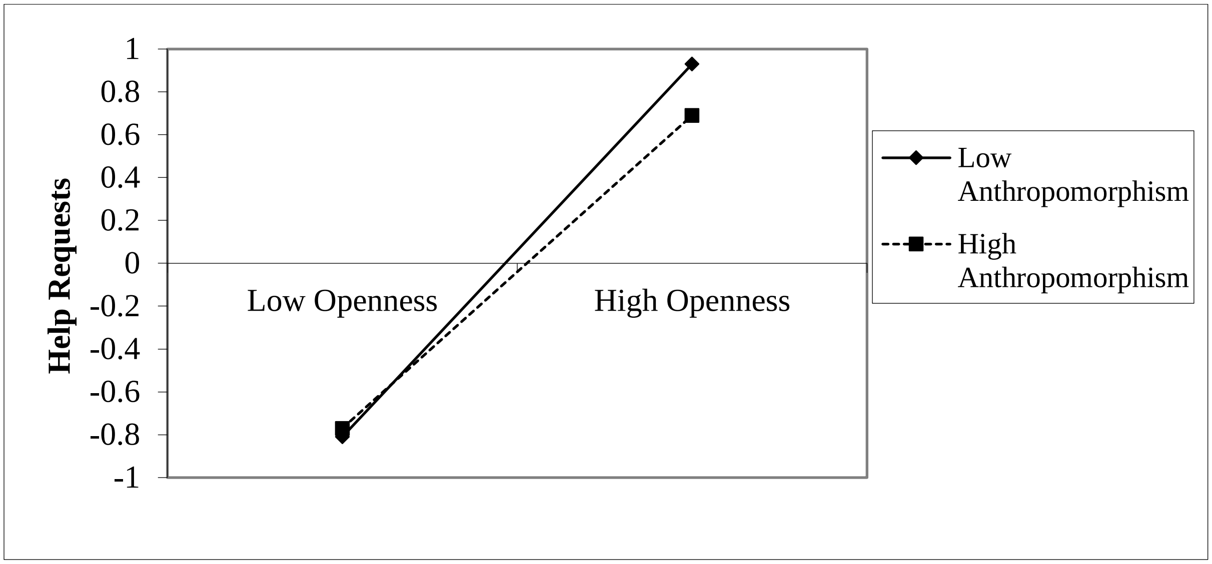
	"This package seems to contain a weapon" signal given		"This package does not seem to contain a weapon" signal given	
	Weapon Present	No Weapon Present	Weapon Present	No Weapon Present
Participant selects "I see a weapon"	**Use** (Appropriate Compliance)	**Misuse** (Inappropriate Compliance)	**Disuse**	**Abuse**
Participant selects "I do not see a weapon"	**Abuse**	**Disuse**	**Misuse** (Inappropriate Reliance)	**Use** (Appropriate Reliance)

Figure 2

Stimuli Presented

Figure 3

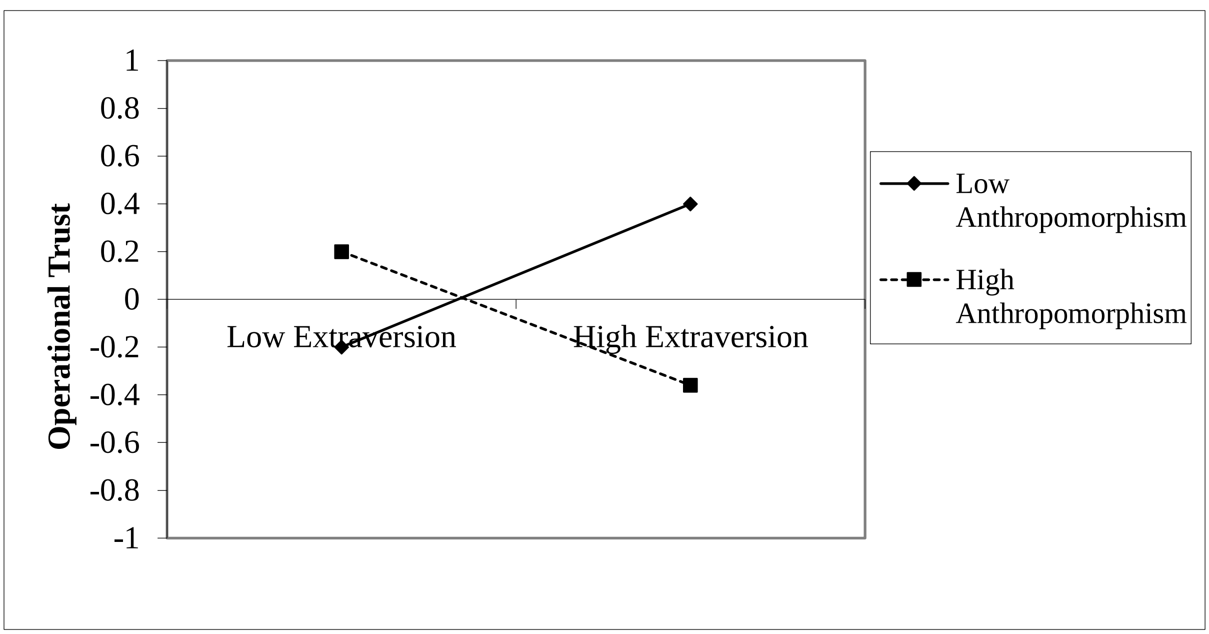
Openness Interaction



Note. All variables standardized.

Figure 4

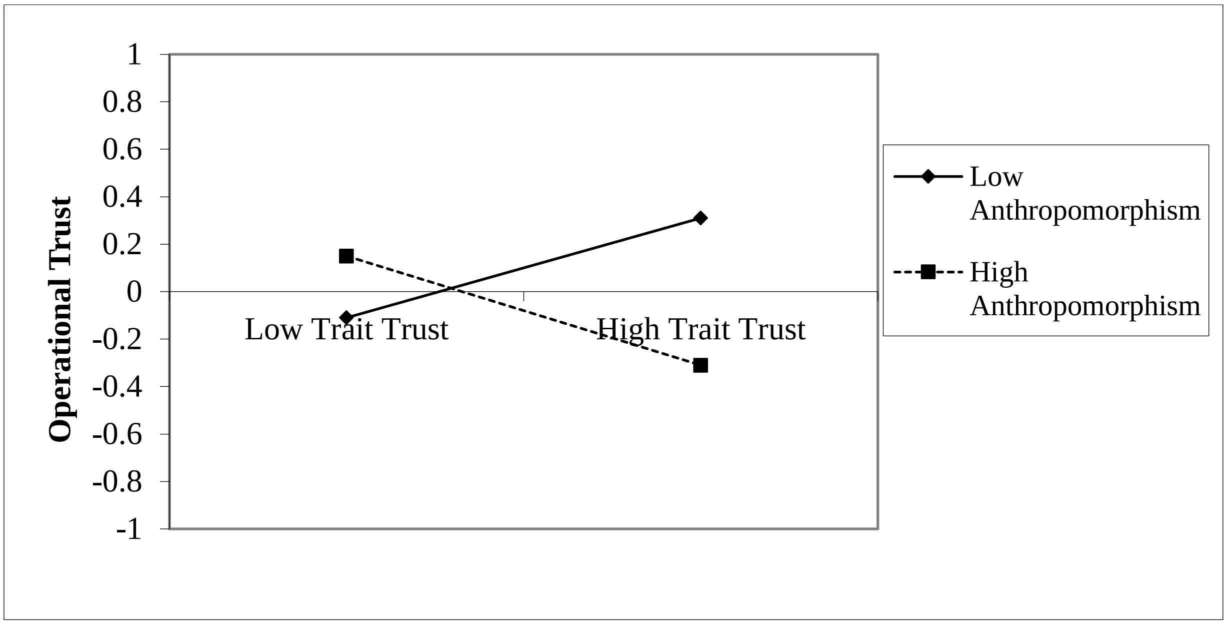
Extraversion Interaction



Note. All variables standardized.

Figure 5

Trait Trust Interaction



Note. All variables standardized.