Achieving Trustworthy Artificial Intelligence: Multi-Source Trust Transfer in Artificial Intelligence-capable Technology

Completed Research Paper

Maximilian Renner
Karlsruhe Institute of Technology
76131 Karlsruhe, Germany
maximilian.renner@kit.edu

Sebastian Lins
Karlsruhe Institute of Technology
76131 Karlsruhe, Germany
lins@kit.edu

Matthias Söllner
University of Kassel
34127 Kassel, Germany
soellner@uni-kassel.de

Scott Thiebes
Karlsruhe Institute of Technology
76131 Karlsruhe, Germany
scott.thiebes@kit.edu

Ali Sunyaev
Karlsruhe Institute of Technology
76131 Karlsruhe, Germany
sunyaev@kit.edu

Abstract

Contemporary research focuses on examining trustworthy AI but neglects to consider trust transfer processes, proposing that users' established trust in a familiar source (e.g., a technology or person) may transfer to a novel target. We argue that such trust transfer processes also occur in the case of novel AI-capable technologies, as they are the result of the convergence of AI with one or more base technologies. We develop a model with a focus on multi-source trust transfer while including the theoretical framework of trust-duality (i.e., trust in providers and trust in technologies) to advance our understanding about trust transfer. A survey among 432 participants confirms that users transfer their trust from known technologies and providers (i.e., vehicle and AI technology) to AI-capable technologies and their providers. The study contributes by providing a novel theoretical perspective on establishing trustworthy AI by validating the importance of the duality of trust.

Keywords: Trustworthy AI, Trust Transfer, Duality of Trust, AI-capable Technologies, Autonomous Vehicles, AI Convergence

Introduction

Artificial intelligence (AI) is one of the driving forces of the so-called "fourth industrial revolution" (Schwab 2017, p. 8). AI's unique aspect is that control is transferred from people to technology, completely changing our previous understanding of people-technology relationships (Schwab 2017). Today, we are particularly witnessing the convergence of AI technology with other common technologies. During convergence, AI technology merges with base technologies while typically taking over the control of users' tasks and enhancing automation (Curran et al. 2010; Duysters and Hagedoorn 1998; Raisch and Krakowski 2021). For example, AI models are nowadays embedded in medical image annotation systems in radiology to provide a higher diagnostic speed (e.g., Miller and Brown 2018). AI convergence may also take the form of intelligent automation such as in the case of autonomous vehicles (AVs) where conventional vehicle technology converges with novel AI technology to automate driving tasks (i.e., self-driving cars; Hengstler et al. 2016; Koester and Salge 2020). Such convergence leads to an ever-increasing number of AI-capable technologies that offer many opportunities to contribute to the well-being of individuals, the economy, and society's progress (Pandl et al. 2020; Yoo et al. 2012).

Providers of AI-capable technologies, however, are faced with users' limited trust in the AI-enhanced functionalities and, thus, users are still hesitant or even relatively hostile to adopt AI-capable technologies (Glikson and Woolley 2020; Hengstler et al. 2016; Liu et al. 2018). Reasons for this hesitation toward AI-capable technologies include risks of infringing individuals' privacy, or the presence of racial bias in widely used AI technology (Thiebes et al. 2020). Besides, users might not be aware of the extent to which technologies converged, and how an AI model's logic provides decisions and functionalities due to these systems' black-box nature (Glikson and Woolley 2020). Indeed, news articles frequently report incidents with AI-capable technologies, such as the case of the world's first pedestrian fatality associated with an AV from Tesla while it was on Autopilot (Forbes 2020). Consequently, providers are looking for powerful means to foster users' trust into their new AI-capable technologies.

Trust has been a central concept in technology acceptance research for decades (Söllner et al. 2016a) and has proven to be a key determinant of individuals' willingness to accept and use a technology because it mitigates uncertainties and risks related to vulnerabilities (Benbasat and Wang 2005; Gefen et al. 2003; Söllner et al. 2012). It is therefore not surprising that the question of how to establish trust in AI(-capable) technologies has become a core discussion in contemporary information systems (IS) research. Several frameworks and guidelines to promote trustworthy AI (TAI) have recently been developed and published by researchers, industry, and policymakers (e.g., Floridi 2019; Independent High-Level Expert Group on Artificial Intelligence 2019; Thiebes et al. 2020). Likewise, research recently examined antecedents of TAI (e.g., explainable AI; Markus et al. 2021) and analyzed the impact of trust on user perceptions (e.g., user satisfaction with AI technology; Shin and Park 2019). While providing valuable contributions, extant research has neglected to consider transfer processes to establish trust in AI-capable technologies. Trust transfer theory proposes that users' trusting beliefs in an already existing and familiar source (e.g., a technology or a person) may transfer to a novel and unknown target (Stewart 2003; Stewart 2006). Such trust transfer typically results if users perceive a strong relationship between a familiar source and an unknown target (Stewart 2003; Stewart 2006). We propose that such trust transfer processes are also likely to occur in the case of novel AI-capable technologies, as they are the result of the convergence of AI with one or more base technologies that, as trust sources, are typically known by users (Glikson and Woolley 2020). For example, if AI technology is embedded in the vehicle technology, leading to AI-capable AVs (Pandl et al. 2020; Shneiderman 2020), users may transfer their established trust in familiar vehicle technologies and supposedly also transfer trust in related AI technologies (e.g., virtual assistants like *Alexa* or *Siri*) to unknown AVs.

However, converging AI into existing technologies challenges existing theoretical assumptions of trust transfer for two reasons. First, AI convergence may evoke a multiple source trust transfer, meaning that users may transfer both their trusting beliefs in a base technology (e.g., vehicle technology) and an AI technology to an AI-capable technology. Although recent research has already validated multi-source trust transfer in related contexts (e.g., Lowry et al. 2014), AI-specifics put in doubt whether a trust transfer from AI technology is achievable. In particular, trust transfer requires users to be familiar with AI technology as a trust source (Stewart 2003; Stewart 2006), yet, users often lack experience or profound knowledge of extant AI technologies due to their novelty and complexity, among others (e.g., Eiband et al. 2021). It remains of high interest to understand if a transfer of trust is occurring and if one of the sources (i.e., base technology or AI technology) has a stronger impact on establishing trust in the target. Second, extant research on trust transfer has mostly focused on either a known technology or provider as a source for trust transfer (e.g., Gong et al. 2020; Stewart 2003). On the contrary, trust research proposes that users' trust typically takes an interwoven dual role: trust in a provider and trust in a technology that need to be considered in parallel (Lansing and Sunyaev 2016; McKnight et al. 2011; Söllner et al. 2016b). The duality of trust is particularly relevant in the context of AI, as users lack understanding of concrete AI technologies, but may be familiar with the AI providers behind them, such as Google, Microsoft, IBM, or Amazon. A differentiated perspective on trust transfer is therefore warranted in the context of AI-capable technologies, which considers trust transfer of both providers and technologies simultaneously to enable comparison. To better understand the trust-building process in AI-capable technologies, we contextualize trust transfer (Hong et al. 2014) and seek to answer the research question (RQ):

RQ: To what extent do users transfer their trust in AI technology and base technologies as known sources from a dual perspective (i.e., trust in technologies and trust in providers) to a converged AI-capable technology as an unknown target?

To answer our research question, we build on trust transfer theory (Stewart 2003; Stewart 2006), and develop a theoretical model with a focus on multi-source trust transfer while including the theoretical framework of trust-duality (i.e., trust in providers and trust in technologies; McKnight et al. 2011). We particularly focus on the context of AVs to examine the trust transfer process within the domain of AI convergence. We tested our theoretical model by conducting an online survey among 432 participants acquired through Amazon Mechanical Turk (MTurk). Our results confirm that users transfer their trust from known vehicle technology and manufacturers and AI technologies and providers to AVs and their providers. The study contributes to IS literature in three key ways. First, we provide a novel theoretical perspective on establishing trust in converging AI-capable technologies by validating the presence of (dual) trust transfer processes in the context of AI. Second, we contribute to trust transfer theory by showing the importance of the duality of trust, proving that trust transfer processes emerge on both a provider and technology trust perspective. Third, we also present a fine-grained perspective of trust in AI technology, measuring it in a traditional way (McKnight et al. 2011), and using AI-technology-specific measurements with the properties: being fair, transparent, accountable and explainable (FATE) (Shin 2020).

This paper proceeds as follows. The next section introduces our understanding of AI-capable technologies and the duality of trust, and describes the theoretical foundations of trust transfer. Afterward, we develop our hypotheses and present our research model. Then, we outline the applied research method to test our hypotheses and present our results. Finally, we outline the implications of our findings, limitations of the study, and opportunities for future research in the discussion section before we briefly conclude this study.

Theoretical Background

AI-Capable Technology and the Case of Autonomous Vehicles

AI-capable technologies refer to technologies that embed AI (e.g., natural language processing, computer vision, or analytics) to enhance their functionalities and take control over users' tasks (Raisch and Krakowski 2021). AI-capable technologies are based on the fact that convergence innovation enables the enhancement of automated intelligence of ubiquitous technologies (Pandl et al. 2020; Yoo et al. 2012). Convergence, in general, describes a phenomenon in which two or more initially separate items merge because of their interplay, their movement toward unity, and their increasing integration with each other (Curran et al. 2010; Duysters and Hagedoorn 1998). This trend impacts several everyday products such as televisions, watches, and vehicles that now have embedded software-based digital capabilities consisting of intelligent automation with sensors, networks, and processors (Yoo et al. 2012). Intelligent automation relies on AI capabilities that enable computers to execute tasks that are easy for people to perform but difficult to describe formally (Pandl et al. 2020). In AI convergence, the impact and embeddedness of AI is typically a step-by-step change, whereas AI is converging more and more with the formally stand-alone base technology. First, AI supports the formally stand-alone base technology while closely collaborating to perform a task (Raisch and Krakowski 2021). Such cases include cognitive systems that analyze medical data to assist physicians in making medical treatment decisions (e.g., Hengstler et al. 2016) or customer service support through AI-based chatbots (e.g., Adam et al. 2020). With the increasing degree of convergence, the impact and embeddedness of AI also increases, while the number of tasks and responsibilities of AI increases and those of humans decreases. AI convergence is thus leading more and more to intelligent automation, with AI taking over the control of formally human tasks (Raisch and Krakowski 2021). Hereby, automation is transforming data to control processes or make decisions (Lee and See 2004) while it is "intelligent", when technology builds its decision-making process and control awareness on inherent AI (Hengstler et al. 2016).

A common research case of intelligent automation and AI convergence are AI-enhanced driving functionalities in AVs (Koester and Salge 2020). In the context of AVs, the step-by-step change of AI convergence is typically divided into six levels of automation (SAE 2018). Level 0 defines a vehicle without any AI capabilities for autonomous driving functionalities and respective intelligent automation. While the level of automation increases, the degree of AI convergence also increases to support autonomous driving functionalities (Hengstler et al. 2016; Shneiderman 2020). For example, in intermediate levels of automation, AI supports the driver with a range of functionalities, such as lane-keeping assistance, or speed control, whereas the drivers continue to be responsible and in control of their vehicles. With higher levels of automation, AI technology takes over more and more actions for the drivers, allowing them to relinquish control of their vehicle to the AI in predefined situations (e.g., on specially upgraded highways). At level 5, AI convergence is most advanced, and AI enhances intelligent automation to match the capabilities of human drivers in most driving scenarios. However, the increasing degree of AI convergence is a double-edged sword and does not only bring advantages but requires more careful consideration of how people establish trust in AI-capable technologies, particularly when AI decisions may impact peoples' well-being as in the case of AV.

Duality of Trust

Trust plays an important role in almost any IS-enabled situation that are characterized by uncertainty or undesirable consequences (McKnight et al. 2011). Nowadays, most IS research adopts a dual perspective on trust, also in the context of TAI (Thiebes et al. 2020). First, trust in people or organizations (Lankton et al. 2015; McKnight et al. 2011), such as trust in a provider (Gefen et al. 2003) or team members (Staples and Webster 2008). Second trust in technology or more specific in an IT artifact (Lankton et al. 2015; McKnight et al. 2011), like a cloud service (Lansing and Sunyaev 2016). Trust in people and trust in technology not only differ on the underlying object but above all on the trusting beliefs (Söllner et al. 2013). Interpersonal trusting beliefs reflect judgments that the other party has appropriate attributes and motives to behave as expected in a risky situation (Mayer et al. 1995), whereas technology-related trust necessarily reflects beliefs about a technology's characteristics rather than of its motives (McKnight et al. 2011; Söllner et al. 2016b). Previous research agrees that individuals can change their expectations about a person's competence (i.e., their ability to do what an individual needs), benevolence (i.e., their care and motivation to act in an individual's interests), and integrity (i.e., their honesty and promise-keeping; McKnight et al. 2002). In contrast, trust in a technology typically refers to the functionality of the technology (i.e., providing features needed to complete a task), its helpfulness (i.e., help functionalities provide necessary advice), and its reliability (i.e., technology will consistently operate properly; McKnight et al. 2011; Thatcher et al. 2011). Nonetheless, these different trust beliefs are strongly interrelated because, for example, a person's competence and a technology's functionality represent individuals' expectations about their capabilities (McKnight et al. 2011). In the case of an AI-capable technology, both lenses on trust play a decisive role, since trust may be established based on users' perceptions toward its technological functionalities and its provider (Thiebes et al. 2020).

Trust Transfer

To understand the trust-building process in AI-capable technologies, we refer to the trust transfer theory that explains the relationship of an already known trusted source and a novel, unknown target (Stewart 2003; Stewart 2006). According to trust transfer theory, four categories are important while understanding the trust transfer mechanism: trust in the source, the source-target relationship, trust in the target, and trusting intention toward the target (Stewart 2003). Toward this end, trust transfer theory indicates that users' trust in a trusted and familiar source can be transferred to a relatively unknown target under the condition that the target is found to be associated with the trusted source (Stewart 2003). While the primary condition is that a user perceives the target to have a strong relationship with the trusted source (Stewart 2003), trust transfer is a fundamental form of trust adjustment between two objects. For example, if users perceive the relationship between a source and a target as close and strong, the transferability of trust is more likely to happen. In contrast, users may not trust the source if the source-target relationship is perceived as weak. Prior research revealed that users might perceive a strong source-target relationship in both a single source context (e.g., from users' trust in web payment services to their trust in mobile payment services; Gong et al. 2020) and a multi-source context (e.g., from users' trust in public administration and the Internet to their trust in the public e-service; Belanche et al. 2014). Yet, research on trust transfer remains scarce (e.g., Gong et al. 2020) and has not been applied to the context of AI-capable technologies. Exploring whether trust transfer also applies to AI contexts provides a fresh perspective on prevalent discussions about TAI and supports researchers in better theorizing the emergence of trust in AI-capable technologies.

Finally, extant research on trust transfer has mostly analyzed trust transfer processes in the context of IT artifacts, such as trust in websites, e-WOM services, and web shopping services (refer to Gong et al. (2020) for a recent and excellent review on trust transfer literature). Research on interpersonal trust transfer, in contrast, is mostly lacking, with exceptions such as with the work of Chen and Shen (2015) that examined trust transfer between trusted members and the community, or Delgado-Ballester and Hernández-Espallardo (2008) that examined how trust in a reputable source brand transfers to online brand extensions. We thus lack a clear understanding whether trust transfer is simultaneously possible at both, a technology and provider level. This understanding is important in the context of AI-capable technologies because users may lack knowledge about AI technologies but may be familiar with their providers.

Research Model

Considering the existing theoretical gaps in trust transfer, we propose to consider multiple trust sources to understand how to establish trust in a converged AI-capable technology as the target, while including a dual perspective of trust (i.e., trust in providers and trust in technologies) to better understand the trust transfer process. We rely on literature from trust transfer and TAI to develop our theoretical model and corresponding hypotheses. To ensure that we consider the converged nature as a multidimensional construct, we include the two sources vehicle and AI while considering trust from a two-sided perspective (Figure 1). We also consider the impact of users' trusting beliefs into the target on users' trusting intention (i.e., using the AV) but do not hypothesize this relationship given its strong support in prior research (e.g., McKnight et al. 2011; McKnight et al. 2002; Stewart 2003).

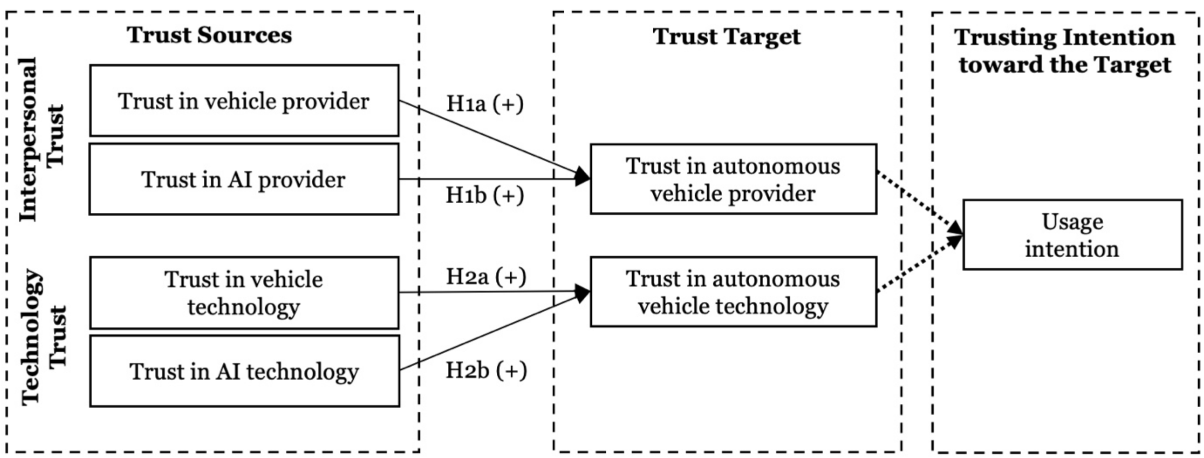

Figure 1. Simplified Research Model

Note: No hypotheses are formulated for dashed relationships.

Trust transfer research proposes that trusting beliefs (i.e., trust in a provider) are transferable from a source to a target (Stewart 2003; Stewart 2006). Hereby, trust transfer is a unique categorization process in that users' trusting beliefs toward a source could be extended to their trusting beliefs toward a target through category-based processing (Stewart 2003; Stewart 2006). Users typically place objects in different categories to classify, interpret, and understand information they receive about these and related objects (Loken et al. 2008). A category is a set of systems, persons, products, or other entities, that appear, to the user, related in some way. For example, users may assign Google, Amazon, and Microsoft to the category *'AI technology provider'*. By grouping objects together that are alike in important respects, users enhance information processing efficiency as well as cognitive stability (Cohen and Basu 1987). A key construct in theoretical accounts of categorization and trust transfer is similarity because it moderates the transfer of cognitive beliefs from one stimulus to another (Loken et al. 2008; Martin and Stewart 2001). If users identify a close similarity between the source and the target, then they are likely to assign the target object in the same category as the source, and transfer knowledge, affect, and intentions to the lesser-known target object.

Building on the trust transfer theory and category-based processing, we argue that users' will put the vehicle provider, AI provider, and AV provider in the same category if users perceive their association. In particular, trust transfer research argues that trust in a target will be influenced by the perceived business relationship between the sources and the target, in other words, the business tie (Lee et al. 2014). For example, if users trust organization A and perceive that organizations A and B are partners, users will trust organization B as well to experience cognitive balance. Although vehicle providers may build the AV themselves, including the intelligent autonomous driving functionalities, more and more providers are taking a different approach in practice and starting collaborative projects with experienced and well-known AI providers. Thus, most AV providers are joint partnership between an existing vehicle provider and AI provider. One example is the association of Daimler with Waymo, which is owned by Alphabet (Daimler 2020). Users may already have encountered other AI technology provided by Alphabet, such as the virtual assistant "Hey Google" or the video recommendations from YouTube. We presume that users put the AV provider and the AI provider and vehicle provider in the same category if they recognize their mutual cooperation to offer the AVs. Since users may be already familiar with AI and vehicle providers, users may swiftly become familiar with the AV provider, especially when the vehicle and AI provider exhibit a strong relationship. We hypothesize:

H1a: Users' trust perception in vehicle providers increases users' trust perception in providers of AV.

H1b: Users' trust perception in AI providers increases users' trust perception in providers of AV.

Likewise, trust transfer research argues that users' trust is transferred based on the perceived technology similarity (Stewart 2003; Stewart 2006). Users may put the source technology and the target technology into one category based on similar technology functionality. Consistent with this, we propose that trust may also be transferred in the case of AI convergence. In the context of AVs, we argue that users will put conventional vehicle technology and AV technology into the same category because it provides similar transportation functionalities. AVs will continue to consist of wheels, doors, and a similar interior, while initially retaining the steering wheel and pedals for the possibility of driver interactions. Nevertheless, more and more new AI-based functionalities will be added, such as voice assistants or the possibility of autonomous driving based on intelligent automation without driver interactions (Koester and Salge 2020). These technological functionalities are similar to previous AI-capable technologies, such as voice assistants in the home environment (e.g., McLean and Osei-Frimpong 2019; Pradhan et al. 2020) or intelligent automated customer service via AI-capable chatbots (e.g., Adam et al. 2020; Zierau et al. 2021). Consequently, users may perceive technology similarities between AVs and their sources (i.e., vehicle technology and AI technology), whereas they may put them into the same category. Thus, trust in AV technology may be transferred from AI technology and vehicle technology. We hypothesize:

H2a: Users' trust perception in vehicle technology increases users' trust perception in AV technology.

H2b: Users' trust perception in AI technology increases users' trust perception in AV technology.

Research Approach

We conducted a cross-sectional survey using online panel data provided by Amazon MTurk to test the research model. Using online panel data has been shown to be suitable for studying trust-related phenomena (e.g., McKnight et al. 2020; Zierau et al. 2021). Research has demonstrated that the results of surveys using MTurk have high reliability and provide high-quality data comparable to student samples or online convenience samples (Buhrmester et al. 2011; Lowry et al. 2016). To design and conduct the survey, we followed established guidelines in IS (Lowry et al. 2016). We restricted potential participants to those with a high reputation (at least 95% approval ratings and at least 5,000 conducted tasks) to ensure sufficiently high data quality (Peer et al. 2013). Also, we restricted participation to US workers to reduce cultural biases, and ensured minimum fair payment of participants (i.e., federal minimum wage of $7.25 per hour). Finally, we embedded attention-check questions to the survey and recorded the time spent on each page to remove responses that had received insufficient attention.

Survey Procedures

We used seven steps to collect survey data. First, we provided a short description of the study's objective, context, and examples of AI technologies (i.e., virtual assistants, recommender systems). Second, we asked subjects to think of a trustworthy AI provider and its provided AI technology they know and like, since familiarity with the source technology is required to enable trust transfer (Stewart 2003; Stewart 2006). We asked subjects to name the AI provider they thought of or select one in a list provided by us (i.e., Microsoft, Apple, Amazon, Google, IBM Watson). We then measured subjects' trust perceptions toward the AI provider and its AI technology. Analogous to this, subjects should next think of a trustworthy vehicle manufacturer and its vehicles they know and like and name it or select one of the provided ones (i.e., Toyota, Ford, VW, Tesla). Note that we have narrowed the scope of vehicles in the survey to "cars" for better subject comprehension. Afterward, we measured subjects' trust in the vehicle manufacturer and its vehicle technology. Third, we added an attention check to control for continued attention and created a washout period between the measurement of our independent and dependent variables by letting subjects read an unrelated text and click on a hidden link (Oppenheimer et al. 2009). Fourth, we introduced subjects into a scenario where they should consider the fictional example that their employer provides them with a company car as part of their salary. They had two options, whereas they could choose from a car with conventional technology and a car with AI-enabled autonomous driving technology. We provided the subjects with brief information to illustrate a strong source-target relationship. In particular, we showed information about the autonomous car provider to illustrate business tie-strength (i.e., a cooperation that has formed between the AI provider and car manufacturer), and the technology to illustrate technical consistency (i.e., autonomous car technology takes over the complete control of the autonomous car when driving on the highway and provides further driver assistant functionalities). Fifth, we measured our dependent variables, starting with subjects' usage intention, followed by the different trust measurements for the AV. Finally, we collected control variables and demographics.

Throughout the survey, we used procedural remedies to reduce the potential bias resulting from common method variance (CMV) (Podsakoff et al. 2003). First, we instructed subjects that answers will be anonymized, that they should take their time to carefully and honestly answer the questions, and that no right or wrong answers exist. Further, we randomized question order, used validated scales from the literature, randomized items, and applied (short) temporal and proximal separation (i.e., different pages) of measurements for independent and dependent variables. Statistical CMV remedies will be explained later.

Survey Measures

We followed methodological recommendations (Straub 1989) and used previously validated scales for measuring the constructs in our survey. For measuring individuals' usage intention we adopted the measurement items from Gong et al. (2020) and Jiang and Benbasat (2007). To measure individuals' trust in technologies, we adopted measures from McKnight et al. (2011), and individuals' trust in providers from Staples and Webster (2008). Since trust in AI has been recently conceptualized as perceiving fair, transparent, accountable and explainable (FATE) AI, we decided to measure trust in AI technology in two ways: first using the conventional trust in technology measures from McKnight et al. (2011) and novel TAI measures from Shin (2021), which comprise items regarding individuals' FATE perceptions. Note that we adapted and rephrased measurement items to fit our context and inserted the name of the proposed or selected car manufacturer and AI provider (i.e., "*I feel comfortable depending on VW for the completion of driving*").

Additionally, we collected control variables to account for potentially confounding influences. First, we controlled for individuals' propensity to trust, particularly toward general technology (McKnight et al. 2011) and humans (Gefen 2000). Second, we added measures to control for the impact of individuals' general attitude toward AI (Schepman and Rodway 2020) on their usage intention. Finally, we controlled for subjects' knowledge of vehicles and AI by adopting the measures from Flynn and Goldsmith (1999).

Descriptive Statistics

We recruited 432 participants, of which we removed 53 responses because 31 participants failed attention checks and 22 participants rushed through the survey. This process resulted in 379 valid responses. This number exceeds the approximate sample size of 198, which we calculated using the tool G*Power (power = 0.95, effect size $f^2 = 0.1$) (Faul et al. 2009) as well as the median sample size of 200 from prior

SEM studies (Kline 2016). More men (32.5% females) participated in our survey, and participants were, on average, 30.4 years old (minimum 23 years, maximum 67 years). Most participants had a high school (18.5%) or undergraduate degree (62.8%, 13.2% graduate degree); have held a driver's license for more than 5 years (86.3%, no driver license 1.3%); had a vehicle which is 3 to 5 years (26.6%) or over 5 years (47.2%) old; and used their vehicle daily (64.4%) or weekly (29.6%). On average, participants indicated that they often interacted with AI technologies (60.7 on a 100-point sliding scale) and rated the realism of the scenario with 82.5 on a 100-point sliding scale.

Data Analysis and Results

We assessed the reliability, convergent validity, and discriminant validity of the constructs (Appendix). All indicators fulfilled the minimum loading requirements (significance and load value) between the indicator and its latent construct, except for four measurement items for the novel FATE AI measures, which we dropped to achieve satisfactory convergent validity (Appendix). The average variance extracted (AVE) was higher than the suggested minimum of .50 (Fornell and Larcker 1981). The composite reliability (CR) values were above .70 demonstrated good internal consistency (Nunnally 1978). Regarding discriminant validity, the square root of each construct's AVE exceeded the inter-construct correlations. In addition, we measured the heterotrait-monotrait (HTMT) ratios of correlations, revealing two issues. First, the HTMT between usage intention and both trust in technology (.890) and trust in provider (.882) slightly exceeds the recommended threshold of .85 (Henseler et al. 2015). We decided to keep usage intention in our model because the Fornell-Larcker Criterion and the less HTMT conservative threshold of .90 are met; usage intention is not the focal construct of our study; and more importantly, because prior theory has already acknowledged a strong relationship between trusting beliefs (i.e., trust in technology and provider) and trusting intentions (i.e., usage intention) (e.g., Gefen 2002; McKnight et al. 2002). Second, the HTMT between trust in technology and trust in provider also exceeded the recommended threshold of .85, which is reasonable given the duality of trust and their high interdependency. To resolve these discriminant validity issues, we decided to calculate three models separately: (1) combining trust in technology and trust in provider as a second-order trust construct and examining a general trust transfer; (2) using trust in provider only and examining whether a provider trust transfer appears; and (3) testing trust in technology only to examine whether we can reveal a technology trust transfer. Finally, we developed a fourth model (4) that measures trust in AI technology using the FATE AI measures instead the conventional trust in technology measures to compare the explanatory power of both measurements.

To test our hypotheses, we used partial least squares structural equation modeling (PLS-SEM) and SmartPLS software, version 3.3.3 (Ringle et al. 2015). The significance of the structural path estimates was assessed using bootstrapping with 5,000 subsamples and with bias-corrected and accelerated confidence intervals (Ringle et al. 2015). We tested the structural model by evaluating the direct effects and the explained variances (R^2). In addition, we controlled for several covariates including individuals' AI and car knowledge, propensity to trust, and general attitude toward AI. We further account for CMV not only ex ante through the careful design of the questionnaire, but also ex post by running a measured latent marker variable (MLMV) test and performing a construct level correction (Chin et al. 2013). For each construct, we added a CMV construct comprising the MLMV items (Appendix), modeled them as impacting each model construct, and compared the bootstrapping results. The differences in the path coefficients between the model constructs were found to be very small (<0.200) (Serrano-Archimi et al. 2018), and we therefore conclude that potential CMV does not pose a significant threat to our results.

The analysis results show that users transfer their trust in both AI and vehicles to AVs (Table 1). Model 1 validates that users' trust in AVs (i.e., combined trust of technology and provider as second-order construct) is positively influenced by users' trust in AI (path coefficient $\beta = .288$; p-value $< .001$; bias-corrected confidence interval [.131, .423]) and users' trust in vehicles ($\beta = .270$; $p < .001$; [.131, .414]). A higher trust in AVs also strongly positively influences users' usage intention ($\beta = .775$; $p < .001$; [.667, .862]).

Path	*path coefficient; p-value; [bias-corrected confidence intervals]*			
	Model 1 General Trust (combined)	Model 2 Trust in Provider	Model 3 Trust in Technology	Model 4 Trust in Technology FATE AI
Trust in AV → Usage Intention	.775; p < .001; [.677, .862]	.674; p < .001; [.588, .752]	.712; p < .001; [.603, .817]	.712; p < .001; [.598, .822]
Trust in AI → Trust in AV	.288; p < .001; [.131, .423]	.345; p < .001; [.199, .480]	.264; p < .001; [.130, .393]	.160; p = .030; [.013, .302]
Trust in Vehicle → Trust in AV	.270; p < .001; [.134, .414]	.270; p < .001; [.141, .394]	.292; p < .001; [.160, .440]	.377; p < .001; [.249, .507]
Controls				
General Attitude AI → Usage Intention	.082; p = .121; [-.016, .189]	.227; p < .001; [.137, .319]	.120; p = .041; [.003, .233]	.120; p = .045; [.000, .238]
AI Knowledge → Usage Intention	.060; p = .041; [.006, .119]	.029; p = .339; [-.028, .091]	.079; p = .018; [.017, .147]	.079; p = .016; [.013, .144]
Vehicle Knowledge → Usage Intention	-.046; p = .207; [-.122, .019]	-.061; p = .078; [-.134, .002]	-.035; p = .376; [-.126, .032]	-.035; p = .383; [-.127, .032]
Propensity to Trust → Trust in AV	.302; p < .001; [.188, .421]	.186; p = .001; [.078, .300]	.346; p < .001; [.239, .459]	.354; p < .001; [.231, .475]
Propensity to Trust → Trust in AI	.421; p < .001; [.321, .501]	.297; p < .001; [.190, .392]	.370; p < .001; [.265, .462]	.433; p < .001; [.334, .521]
Propensity to Trust → Trust in Vehicle	.303; p < .001; [.210, .383]	.269; p < .001; [.170, .365]	.242; p < .001; [.154, .328]	.242; p < .001; [.150, .331]

Table 1. PLS Results for Each Model

Note: grey-filled cells indicate at least p < .05.

Focusing on a provider trust transfer in Model 2 reveals that users' trust in a vehicle provider positively impacts users' trust in the AV provider (β = .270; p < .001; [.141, .394]), **supporting H1a**. Similarly, users' trust in an AI provider positively impacts users' trust in the AV provider (β = .345; p < .001; [.199, .480]), **supporting H1b**. We also identify a similar trust transfer effect when focusing on users' trust in technology in Model 3. Users' trust in the technology of AVs is positively impacted by users' trust in vehicle technology (β = .292; p < .001; [.160, .440]), **supporting H2a**; and users' trust in AI technology (β = .264; p < .001; [.130, .393]), **supporting H2b**. Model 4 substantiates these effects by showing that users' trust in AI technology measured by using the FATE operationalization of TAI (Shin 2020) also positively impacts users' trust in AV technology (β = .160; p = .030; [.013, .302]).

Discussion

Principal Findings

We aimed to better understand the extent to which trust may be transferable in an AI convergence use case. Our study yields three key findings, which are summarized in Table 2.

Previous research gaps	Key findings
Prior research shows that trust in a new and unknown target may be transferred in a multiple source context (e.g., Lowry et al. 2014). However, we know little about how trust is established in the context of AI convergence and whether trust transfer from AI technology as a source is achievable. More research is needed to understand whether trust transfer theory is also suitable to understand the establishment of trust in AI-capable technologies.	Multiple sources need to be considered when transferring trust in AI-capable technologies and their providers. In the context of AVs, users perceive already known AI technologies and the conventional vehicle technologies and their respective providers as trust sources.
Extant research already shows that trust in a new and unknown target may be transferred from different trust sources. However, research considered either a known technology or provider as source (e.g., Gong et al. 2020; Stewart 2003). More research is needed to determine if a dual trust perspective considering both, providers and technologies, is also applicable to the same object of trust transfer.	Trust transfer between trust in vehicles and trust in AI as sources toward trust in AVs as a target occurs from a dual trust perspective (i.e., trust in providers and trust in technologies).
Trust in technologies is commonly measured with functionality, helpfulness, and reliability (McKnight et al. 2011). However, current research started to use AI-specific trust measurements (Shin 2020). More research is needed to understand how to measure TAI.	Both types of trust measurement were applicable, while the conventional trust items achieved better internal validity than AI-specific items.

Table 2. Summary of the Key Findings

First, our findings show that trust is transferable from AI and vehicles to AVs on a provider and technology level. While we were at the beginning of our study uncertain whether users also transfer trust into AI providers and technologies to target objects due to AI-peculiarities, our findings support that trust in related AI technologies also may become an important trust source. Comparing the effect-sizes of the models 1,2 and 3 reveals that the impact of AI as trust source is equally strong compared to the impact of trust in vehicles as base technology and further trust source (i.e., small effects, $d <= 0.2$). Our results thus do not reveal a predominance of one of the trust sources but rather confirm that researchers should consider both when theorizing about TAI. Furthermore, our research shows that users perceive the convergence of AI technology and other base technologies as sources that need to be considered when establishing trust in an AI-capable technology. Second, by distinguishing two types of trust (i.e., trust in providers and trust in technologies) and investigating their impact on AI and vehicles as sources and AVs as a target, we were able to validate the importance of a dual trust perspective for trust transfer processes. Thereby, we find that both trust perspectives have an impact on establishing trust and are particularly relevant in the context of AI convergence. Finally, we also compared conventional trust in technology measures with novel AI-specific items to consider recent developments in the field of TAI. In general, both measures were applicable to measure trust in AI technology. Nevertheless, we needed to drop four measurement items for the novel FATE AI measures to achieve satisfactory convergent validity during our CFA (Appendix). Interestingly, these four items relate to the accountability (i.e., items TF3 & TF4) and transparency (i.e., items TF5 & TF6) dimensions of TAI, whereas we kept items for measuring fairness and explainability of AI technology. One explanation might be that survey subjects found it challenging to elaborate on accountability and transparency of prevalent AI technologies given their complexity and black-box nature.

Theoretical and Practical Contributions

Our research findings have several theoretical contributions. First, our research provides a novel perspective on how to establish trust in AI-capable technologies. By contextualizing trust transfer theory (Stewart 2003; Stewart 2006), we find that trust in novel AI-capable technologies is transferable from familiar and known AI and base technology sources. Thus, we provide theoretical arguments for why it is important to consider the base technology and the already familiar AI technologies when theorizing how trust may be established in a new converged AI technology. So far, IS research on trust-building has considered AI technologies as separate and self-contained objects, neglecting the source objects and thus trust transfer (e.g., Shin 2020). In contrast, our results show that if users recognize the functionalities or providers of an AI technology, they may also transfer their beliefs to new unknown AI-capable technologies once they put them in the same category. Toward this end, we contribute to research by showing that trust transfer also occurs, and for establishing trust in an AI-capable technology, its converging sources must also be considered. Second, we illustrated the impact of a dual trust perspective on trust transfer. By considering the recommended duality of trust (i.e., trust in providers and trust in technologies; Lankton et al. 2015; Lansing and Sunyaev 2016; McKnight et al. 2011), we also provide a more nuanced view on trust transfer processes and thereby inform trust transfer theory. To the best of our knowledge, the duality of trust has not been explored in previous trust transfer research. While researchers have suggested that different trust perspectives can have an impact on a target as separate sources, the duality of trust has not been applied to the same source (Gong et al. 2020). Third, we provide evidence about different TAI measurements. In doing so, we have shown that both conventional measurement (McKnight et al. 2011) and new AI-specific measurement (Shin 2020) are applicable for understanding TAI from a technology perspective. Nevertheless, our study highlights first criticism regarding the validity of AI-specific FATE measurements because individuals may not be able to full elaborate on each FATE dimension.

Our study yields practical implications for how to establish trust in AI-capable technologies. First, we show that users may perceive both the technology and the provider of an AI-enabled technology as important trust sources. We propose that, especially in AI convergence, a provider perspective is interesting because users may better know and be familiar with providers than the technological functionalities. Thus, providers of AI-capable technologies should not only aim to ensure that users establish trust in the technological functionality itself and perceive the added value of, for example, intelligent automation in vehicles. The individual technology providers (i.e., vehicle and AI providers) and the underlying association of them (i.e., strong business tie-strength) must also be trustworthy and thus must be taken into account. Providers of AVs should also keep in mind that not only the vehicle provider, but also the AI provider may be required to be clearly perceivable to the users. Second, we have shown that users perceive an AV as the result of the convergence of AI and vehicle technology. We demonstrated that trust in AVs is established by trust in the conventional vehicle and trust in AI. Providers of AVs should bear in mind that both sources should be considered to understand when and how users may establish trust in an AVs with intelligent automation functionalities.

Limitations and Future Research

Our study has some limitations that open avenues for future research. First, we collected data during the COVID-19 pandemic. The pandemic substantially affected people's private and professional life and might have affected our data collection as well (Prommegger et al. 2021). Second, we collected data for dependent and independent variables simultaneously, which could cause CMV in our data. While we carefully designed our study to minimize the risk, future research could separate data collection over time or use other data types (e.g., sales data). Third, our study uses the online platform MTurk for the selection of study participants. While prior research acknowledges MTurk's suitability for behavioral studies (e.g., Lowry et al. 2016), the generalizability of the results should be taken with caution. AI-capable technology usage particularly depends on the users and their characteristics and attitudes. The issue, however, with the participants at MTurk is that they represent a younger (in our case an average age of 30.4 years) and more educated population (in our case, for example, 62.8% with an undergraduate degree). Thus, future research should employ additional means of data collection including a more diverse population. For instance, engaging multiple online panel providers, conducting behavioral experiments, or investigate related scenarios (e.g., vehicle voice assistants) could help to triangulate insights.

With this study we wanted to create opportunities for behavioral research to yield fresh insight into TAI. Future research may try to better understand under which circumstances trust is transferred in the context of AI-capable technologies. Previous trust transfer research has shown that the source-target relationship must be strong and perceived in order for trust to be transferred (Stewart 2003). Although our study shows that trust in AI convergence is transferable, we did not consider the source-target relationship and did not consider any moderating or mediating effects. Further research may also consider other AI convergence use cases and domains besides the context of the AVs to evaluate if trust is transferable in AI-capable technologies from established trusting beliefs in already known and familiar sources.

Conclusion

This study aimed to understand whether trust transfer is also occurring in an AI convergence context. We find that trust transfer may explain how users establish trust in new AI-capable technologies while considering the underlying base technologies and already known and familiar AI technologies. We suggested that a multiple source perspective and a dual trust perspective (i.e., trust in providers and trust in technologies) must be considered to fully understand trust transfer processes. Regarding our use case AV, we confirm that AI and vehicles are perceived as sources of trust transfer. Trust is transferred on both the provider and the technology level, whereby the participants also perceive the technological and provider convergence. Thus, we not only show that trust transfer is relevant for establishing trust in AI-capable technologies, but also that a dual trust perspective is needed. Although trust may be transferred to new AI-capable technologies through technological functionalities, as in the case of familiar intelligent automated solutions (e.g., AI-based chatbots), the provider-side should not be neglected. Users may not always understand the AI functionalities, which is why the measurement of trust in technology with AI-specific items may have performed worse than with traditional items. However, the AI providers may be known to users and thus provide a solution for building trust even without technological understanding. This study provides promising insights into why further research is needed to understand the processes of why and how trust is transferable in AI convergence.

References

Adam, M., Wessel, M., and Benlian, A. 2020. "AI-Based Chatbots in Customer Service and Their Effects on User Compliance," *Electronic Markets* (31:2), pp. 427–445.

Belanche, D., Casaló, L. V., Flavián, C., and Schepers, J. 2014. "Trust Transfer in the Continued Usage of Public E-Services," *Information & Management* (51:6), pp. 627-640.

Benbasat, I., and Wang, W. 2005. "Trust in and Adoption of Online Recommendation Agents," *Journal of the Association for Information Systems* (6:3), pp. 72-101.

Buhrmester, M., Kwang, T., and Gosling, S. D. 2011. "Amazon's Mechanical Turk: A New Source of Inexpensive, yet High-Quality, Data?," *Perspectives on Psychological Science* (6:1), pp. 3-5.

Chen, J., and Shen, X.-L. 2015. "Consumers' Decisions in Social Commerce Context: An Empirical Investigation," *Decision Support Systems* (79), pp. 55-64.

Chin, W. W., Thatcher, J. B., Wright, R. T., and Steel, D. 2013. "Controlling for Common Method Variance in PLS Analysis: The Measured Latent Marker Variable Approach," in *New Perspectives in Partial Least Squares and Related Methods,* H. Abdi, W. W. Chin, V. Esposito Vinzi, G. Russolillo and L. Trinchera (eds.), New York: Springer, pp. 231-239.

Cohen, J. B., and Basu, K. 1987. "Alternative Models of Categorization: Toward a Contingent Processing Framework," *Journal of Consumer Research* (13:4), pp. 455-472.

Curran, C.-S., Bröring, S., and Leker, J. 2010. "Anticipating Converging Industries Using Publicly Available Data," *Technological Forecasting and Social Change* (77:3), pp. 385-395.

Daimler. 2020. "Strategic Partnership. Daimler Trucks and Waymo," (available online at: https://www.daimler.com/investors/reports-news/financial-news/20201027-dt-and-waymo.html, last access September 1, 2021)

Delgado-Ballester, E., and Hernández-Espallardo, M. 2008. "Effect of Brand Associations on Consumer Reactions to Unknown On-Line Brands," *International Journal of Electronic Commerce* (12:3), pp. 81-113.

Duysters, G., and Hagedoorn, J. 1998. "Technological Convergence in the IT Industry: The Role of Strategic Technology Alliances and Technological Competencies," *International Journal of the Economics of Business* (5:3), pp. 355-368.

Eiband, M., Buschek, D., and Hussmann, H. 2021. "How to Support Users in Understanding Intelligent Systems? Structuring the Discussion," in *Proceedings of the 26th International Conference on Intelligent User Interfaces*, College Station, TX, USA pp. 120–132.

Faul, F., Erdfelder, E., Buchner, A., and Lang, A.-G. 2009. "Statistical Power Analyses Using G*Power 3.1: Tests for Correlation and Regression Analyses," *Behavior Research Methods* (41:4), pp. 1149-1160.

Floridi, L. 2019. "Establishing the Rules for Building Trustworthy AI," *Nature Machine Intelligence* (1:6), pp. 261-262.

Flynn, L. R., and Goldsmith, R. E. 1999. "A Short, Reliable Measure of Subjective Knowledge," *Journal of Business Research* (46:1), pp. 57-66.

Forbes. 2020. "Tesla Lawsuit over Autopilot-Engaged Pedestrian Death Could Disrupt Automated Driving Progress," (available online at: https://www.forbes.com/sites/lanceeliot/2020/05/16/lawsuit-against-tesla-for-autopilot-engaged-pedestrian-death-could-disrupt-full-self-driving-progress/?sh=7af1a10371f4 last access September 1, 2021)

Fornell, C., and Larcker, D. F. 1981. "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research* (18:1), pp. 39-50.

Gefen, D. 2000. "E-Commerce: The Role of Familiarity and Trust," *Omega* (28:6), pp. 725-737.

Gefen, D. 2002. "Customer Loyalty in E-Commerce," *Journal of the Association for Information Systems* (3:1), pp. 27-51.

Gefen, D., Karahanna, E. W., and Straub, D. 2003. "Trust and TAM in Online Shopping: An Integrated Model," *MIS Quarterly* (27:1), pp. 51-90.

Glikson, E., and Woolley, A. W. 2020. "Human Trust in Artificial Intelligence: Review of Empirical Research," *Academy of Management Annals* (14:2), pp. 627-660.

Gong, X., Zhang, K. Z. K., Chen, C., Cheung, C. M. K., and Lee, M. K. O. 2020. "What Drives Trust Transfer from Web to Mobile Payment Services? The Dual Effects of Perceived Entitativity," *Information & Management* (57:7), pp. 1-11.

Hengstler, M., Enkel, E., and Duelli, S. 2016. "Applied Artificial Intelligence and Trust—the Case of Autonomous Vehicles and Medical Assistance Devices," *Technological Forecasting and Social Change* (105), pp. 105-120.

Henseler, J., Ringle, C. M., and Sarstedt, M. 2015. "A New Criterion for Assessing Discriminant Validity in Variance-Based Structural Equation Modeling," *Journal of the Academy of Marketing Science* (43:1), pp. 115-135.

Hong, W., Chan, F. K. Y., Thong, J. Y. L., Chasalow, L. C., and Dhillon, G. 2014. "A Framework and Guidelines for Context-Specific Theorizing in Information Systems Research," *Information Systems Research* (25:1), pp. 111-136.

Independent High-Level Expert Group on Artificial Intelligence. 2019. "Ethics Guidelines for Trustworthy AI," European Commission, Brussels.

Jiang, Z., and Benbasat, I. 2007. "Investigating the Influence of the Functional Mechanisms of Online Product Presentations," *Information Systems Research* (18:4), pp. 454-470.

Koester, N., and Salge, T. O. 2020. "Building Trust in Intelligent Automation: Insights into Structural Assurance Mechanisms for Autonomous Vehicles," in *Proceedings of the 41st International Conference on Information Systems (ICIS)*, India pp. 2143-2160.

Lankton, N. K., McKnight, D. H., and Tripp, J. 2015. "Technology, Humanness, and Trust: Rethinking Trust in Technology," *Journal of the Association for Information Systems* (16:10), pp. 880-918.

Lansing, J., and Sunyaev, A. 2016. "Trust in Cloud Computing: Conceptual Typology and Trust-Building Antecedents," *ACM SIGMIS Database* (47:2), pp. 58-96.

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

Lee, K. C., Lee, S., and Hwang, Y. 2014. "The Impact of Hyperlink Affordance, Psychological Reactance, and Perceived Business Tie on Trust Transfer," *Computers in Human Behavior* (30), pp. 110-120.

Liu, P., Yang, R., and Xu, Z. 2018. "Public Acceptance of Fully Automated Driving: Effects of Social Trust and Risk/Benefit Perceptions," *Risk Analysis* (39:2), pp. 326-341.

Loken, B., Barsalou, L. W., and Joiner, C. 2008. "Categorization Theory and Research in Consumer Psychology: Category Representation and Category-Based Inference," in *Handbook of Consumer Psychology,* C. P. Haugtvedt, P. M. Herr and F.R. Kardes (eds.), Mahwah, NJ: Lawrence Erlbaum Associates, pp. 133-163.

Lowry, P. B., D'Arcy, J., Hammer, B., and Moody, G. D. 2016. "'Cargo Cult' Science in Traditional Organization and Information Systems Survey Research: A Case for Using Nontraditional Methods of Data Collection, Including Mechanical Turk and Online Panels," *The Journal of Strategic Information Systems* (25:3), pp. 232-240.

Lowry, P. B., Twyman, N. W., Pickard, M., Jenkins, J. L., and Bui, Q. N. 2014. "Proposing the Affect-Trust Infusion Model (ATIM) to Explain and Predict the Influence of High and Low Affect Infusion on Web Vendor Trust," *Information & Management* (51:5), pp. 579-594.

Markus, A. F., Kors, J. A., and Rijnbeek, P. R. 2021. “The Role of Explainability in Creating Trustworthy Artificial Intelligence for Health Care: A Comprehensive Survey of the Terminology, Design Choices, and Evaluation Strategies,” *Journal of Biomedical Informatics* (113), pp. 1-11.

Martin, I. M., and Stewart, D. W. 2001. “The Differential Impact of Goal Congruency on Attitudes, Intentions, and the Transfer of Brand Equity,” *Journal of Marketing Research* (38:4), pp. 471-484.

Mayer, R. C., Davis, J. H., and Schoorman, F. D. 1995. “An Integrative Model of Organizational Trust,” *The Academy of Management Review* (20:3), pp. 709-734.

McKnight, D. H., Carter, M., Thatcher, J. B., and Clay, P. F. 2011. “Trust in a Specific Technology,” *ACM Transactions on Management Information Systems* (2:2), pp. 1-25.

McKnight, D. H., Choudhury, V., and Kacmar, C. 2002. “Developing and Validating Trust Measures for E-Commerce: An Integrative Typology,” *Information Systems Research* (13:3), pp. 334-359.

McKnight, D. H., Liu, P., and Pentland, B. T. 2020. “Trust Change in Information Technology Products,” *Journal of Management Information Systems* (37:4), pp. 1015-1046.

McLean, G., and Osei-Frimpong, K. 2019. “Hey Alexa … Examine the Variables Influencing the Use of Artificial Intelligent in-Home Voice Assistants,” *Computers in Human Behavior* (99), pp. 28-37.

Miller, D. D., and Brown, E. W. 2018. “Artificial Intelligence in Medical Practice: The Question to the Answer?,” *The American Journal of Medicine* (131:2), pp. 129-133.

Nunnally, J. C. 1978. “Psychometric Theory (2nd Ed.),” New York, NY: McGraw-Hill College, pp. 1-701.

Oppenheimer, D. M., Meyvis, T., and Davidenko, N. 2009. “Instructional Manipulation Checks: Detecting Satisficing to Increase Statistical Power,” *Journal of Experimental Social Psychology* (45:4), pp. 867-872.

Pandl, K. D., Thiebes, S., Schmidt-Kraepelin, M., and Sunyaev, A. 2020. “On the Convergence of Artificial Intelligence and Distributed Ledger Technology: A Scoping Review and Future Research Agenda,” *IEEE Access* (8), pp. 57075-57095.

Peer, E., Vosgerau, J., and Acquisti, A. 2013. “Reputation as a Sufficient Condition for Data Quality on Amazon Mechanical Turk,” *Behavior Research Methods* (46:4), pp. 1023-1031.

Podsakoff, P., MacKenzie, S., Lee, J.-Y., and Podsakoff, N. 2003. “Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies,” *The Journal of Applied Psychology* (88:5), pp. 879-903.

Pradhan, A., Lazar, A., and Findlater, L. 2020. “Use of Intelligent Voice Assistants by Older Adults with Low Technology Use,” *ACM Transactions on Computer-Human Interaction* (27:4), pp. 1-27.

Prommegger, B., Bennett Thatcher, J. B., Wiesche, M., and Krcmar, H. 2021. “When Your Data Has COVID-19: How the Changing Context Disrupts Data Collection and What to Do About It,” *European Journal of Information Systems* (30:1), pp. 100-118.

Raisch, S., and Krakowski, S. 2021. “Artificial Intelligence and Management: The Automation–Augmentation Paradox,” *Academy of Management Review* (46:1), pp. 192-210.

Ringle, C., Wende, S., and Becker, J.-M. 2015. “SmartPLS 3 (Version 3),” SmartPLS GmbH, Boenningstedt, Germany. (www.smartpls.com).

SAE. 2018. “Taxonomy and Definitions for Terms Related to Driving Automation Systems for on-Road Motor Vehicles,” SAE International (J3016).

Schepman, A., and Rodway, P. 2020. “Initial Validation of the General Attitudes Towards Artificial Intelligence Scale,” *Computers in Human Behavior Reports* (1), pp. 1-13.

Schwab, K. 2017. “The Fourth Industrial Revolution,” New York, NY: Portfolio Penguin, pp. 1-184.

Serrano-Archimi, C., Reynaud, E., Yasin, H. M., and Bhatti, Z. A. 2018. “How Perceived Corporate Social Responsibility Affects Employee Cynicism: The Mediating Role of Organizational Trust,” *Journal of Business Ethics* (151:4), pp. 907-921.

Shin, D. 2020. “User Perceptions of Algorithmic Decisions in the Personalized AI System:Perceptual Evaluation of Fairness, Accountability, Transparency, and Explainability,” *Journal of Broadcasting & Electronic Media* (64:4), pp. 541-565.

Shin, D. 2021. “The Effects of Explainability and Causability on Perception, Trust, and Acceptance: Implications for Explainable AI,” *International Journal of Human-Computer Studies* (146), pp. 1-40.

Shin, D., and Park, Y. J. 2019. “Role of Fairness, Accountability, and Transparency in Algorithmic Affordance,” *Computers in Human Behavior* (98), pp. 277-284.

Shneiderman, B. 2020. “Bridging the Gap between Ethics and Practice: Guidelines for Reliable, Safe, and Trustworthy Human-Centered AI Systems,” *ACM Transactions on Interactive Intelligent Systems* (10:4), pp. 1-31.

Söllner, M., Benbasat, I., Leinmeister, J. M., and Pavlou, P. A. 2016a. "Trust," in *MIS Quaterly Research Curations,* A. Bush and A. Rai (eds.).
Söllner, M., Hoffmann, A., Hoffmann, H., Wacker, A., and Leimeister, J. M. 2012. "Understanding the Formation of Trust in IT Artifacts," in *Proceedings of the 34th International Conference on Information Systems (ICIS),* Orlando, FL, USA.
Söllner, M., Hoffmann, A., and Leimeister, J. M. 2016b. "Why Different Trust Relationships Matter for Information Systems Users," *European Journal of Information Systems* (25:3), pp. 274-287.
Söllner, M., Pavlou, P., and Leimeister, J. M. 2013. "Understanding Trust in IT Artifacts - a New Conceptual Approach," in *Proceedings of the Academy of Management Annual Meeting* Orlando, FL, USA.
Staples, D. S., and Webster, J. 2008. "Exploring the Effects of Trust, Task Interdependence and Virtualness on Knowledge Sharing in Teams," *Information Systems Journal* (18:6), pp. 617-640.
Stewart, K. J. 2003. "Trust Transfer on the World Wide Web," *Organization Science* (14:1), pp. 5-17.
Stewart, K. J. 2006. "How Hypertext Links Influence Consumer Perceptions to Build and Degrade Trust Online," *Journal of Management Information Systems* (23:1), pp. 183-210.
Straub, D. W. 1989. "Validating Instruments in MIS Research," *MIS Quarterly* (13:2), pp. 147-169.
Thatcher, J. B., McKnight, D. H., White Baker, E., Arsal, R., and Roberts, N. H. 2011. "The Role of Trust in Postadoption IT Exploration: An Empirical Examination of Knowledge Management Systems," *Engineering Management, IEEE Transactions on* (58:1), pp. 56-70.
Thiebes, S., Lins, S., and Sunyaev, A. 2020. "Trustworthy Artificial Intelligence," *Electronic Markets* (31:2), pp. 447–464.
Yoo, Y., Boland, R. j., Lyytinen, K., and Majchrzak, A. 2012. "Organizing for Innovation in the Digitized World," *Organization Science* (23:5), pp. 1398-1408.
Zierau, N., Flock, K., Janson, A., Söllner, M., and Leimeister, J. M. 2021. "The Influence of AI-Based Chatbots and Their Design on Users Trust and Information Sharing in Online Loan Applications," in *Proceedings of the 54th Hawaii International Conference on System Sciences (HICSS)*, Koloa, Hawaii, USA, pp. 5483-5492.

Appendix: Measurement

Label	Item	Mean (SD)	Loading
[Vehicle / AI] Knowledge (Flynn and Goldsmith 1999) (α = [.95 / .95], CR = [.96 / .96], AVE = [.88. / .88])			
VK1	Among my circle of friends, I'm one of the "experts" on [cars / AI].	[43.43 (30.77) / 45.40 (28.58)]	[.946 /.923]
VK2	I do feel very knowledgeable about [cars / AI].	[47.26 (29.40) / 46.50 (27.04)]	[.864 /.934]
VK3	Compared to most other people, I know much about [cars / AI].	[46.20 (29.54) / 49.10 (26.73)]	[.973 / .941]
VK4	I know pretty much about [cars / AI].	[47.10 (28.58) / 47.18 (26.96)]	[.965 /.958]

Label	Item	Mean (SD)	Loading
General Attitude AI toward AI (Schepman and Rodway 2020)(α = .92, CR = .93, AVE = .65).			
GAI1	There are many beneficial applications of AI technology.	81.13 (19.79)	.873
GAI2	I am impressed by what AI technology can do.	80.73 (19.97)	.856
GAI3	AI technology can have positive impacts on people's wellbeing.	79.72 (20.87)	.874
GAI4	AI technology is exciting.	77.69 (22.98)	.866
GAI5	AI technology can provide new economic opportunities for this country.	77.70 (21.77)	.823
GAI6	AI-based systems can perform better than humans.	67.20 (22.60)	.678
GAI7	Much of society will benefit from a future full of AI technology.	75.48 (22.88)	.886
GAI8	For routine transactions, I would rather interact with an AI-based system than with a human.	63.19 (27.44)	.557
Trust Propensity – Technology (McKnight et al. 2011) (α = .89, CR = .93, AVE = .83).			
TPT1	I generally give a technology the benefit of the doubt when I first use it.	69.78 (24.03)	.920
TPT2	I usually trust a technology until it gives me a reason not to trust it.	68.98 (25.34)	.926
TPT3	My typical approach is to trust new technologies until they prove to me that I shouldn't trust them.	66.42 (25.37)	.890
Trust Propensity – Humans (Gefen 2000)*(α = .93, CR = .95, AVE = .83).*			
TPH1	I feel that people are generally reliable.	63.58 (25.38)	.932
TPH2	I generally have faith in humanity.	65.16 (25.57)	.920
TPH3	I generally trust other people unless they give me reason not to.	65.48 (26.75)	.894
TPH4	I tend to count upon other people.	61.13 (26.70)	.914
Usage Intention (Gong et al. 2020; Jiang and Benbasat 2007)(α = .94, CR = .96, AVE = .89).			
UI1	Given that I have access to the autonomous car, I predict that I would choose it.	72.38 (28.46)	.949
UI2	I am very likely to choose the autonomous car with the information it needs to better serve my needs.	73.54 (28.27)	.955
UI3	I would recommend to a friend the use of the autonomous car.	68.45 (28.19)	.932
Trust in Provider of [Vehicle / AI / Autonomous Vehicle] (Staples and Webster 2008) (α = [.83 / .86 / .91], CR = [.89 / .91 / .94], AVE = [.74 / .79 / .85].			
TP1	I am comfortable letting [Vehicle Manufacturer / AI Provider / *Autonomous Vehicle*] take responsibility for tasks which are critical to [Vehicle / AI / Autonomous Vehicle Technology] even when I cannot control them.	[76.27 (22.56) / 68.51 (23.43) / 65.23 (27.74)]	[.863 /.872 /.911]
TP2	I feel comfortable depending on [Vehicle Manufacturer / AI Provider / *Autonomous Vehicle Provider*] for the completion of AI-supported tasks.	[82.60 (19.30)/ 73.08 (22.48) / 70.71 (25.99)]	[.883 /.898 /.936]
TP3	Overall, I feel that I can trust [Vehicle Manufacturer / AI Provider / *Autonomous Vehicle Provider*] completely.	[77.83 (21.75) / 67.33 (26.30) / 70.47 (26.50)]	[.850 /.897 /.919]

Label	Item	Mean (SD)	Loading
Trust in Technology [Vehicle / AI / Autonomous Vehicle] (McKnight et al. 2011) *(α = [.94 / .94 / .95], CR = [.95 / .95 / .96], AVE = [.73 / .73 / .78]. The [Vehicle / AI / Autonomous Vehicle]-technology…*			
TT1	… is a very reliable technology.	[82.64 (17.75) / 78.93 (18.38) / 74.12 (23.23)]	[.877 /.866 /.901]
TT2	… is extremely dependable.	[81.78 (18.87)/ 75.66 (19.62) /73.07 (23.93)]	[.896 /.882 /.920]
TT3	… does not fail me.	[79.42 (21.15) / 72.45 (22.11) / 71.14 (24.72)]	[.852 /.848 / .903]
TT4	… does not malfunction for me.	[78.53 (21.86) / 71.87 (23.35) / 71.08 (24.34)]	[.784 / .819 / .896]
TT5	… has the ability to do what I want it to do.	[85.73 (17.25) / 78.05 (19.70) / 77.86 (22.56)]	[.877 / .843 / .845]
TT6	… has the features required to fulfill my needs.	[86.10 (19.00) / 77.54 (20.03) / 82.28 (20.57)]	[.846 / .876 / .800]
TT7	… has the functionality I need.	[84.62 (18.00) / 77.96 (20.41) / 77.09 (23.32)]	[.867 / .870 / .909]
Trust in AI Technology FATE (Shin 2021)(α = .74, CR = .84, AVE = .57)			
TF1	In my opinion, [AI-technology] has no favoritism and does not discriminate against people.	76.63 (24.03)	.834
TF2	I believe [AI-technology] follows due process of impartiality with no prejudice.	76.51 (23.06)	.860
TF3	I think that [AI-technology] has a person in charge who is accountable for its adverse individual or societal effects in a timely fashion.	62.10 (27.19)	*
TF4	I think [AI-technology] is designed to enable third parties to examine and review the behavior of an algorithm.	59.92 (28.69)	*
TF5	I think that the evaluation and the criteria of [AI-technology] is publicly released and understandable to people.	51.64 (29.91)	*
TF6	In my opinion, [AI-technology] lets people know how well internal states of algorithms are understood from knowledge of its external outputs.	53.12 (29.16)	*
TF7	I think the [AI-technology] is interpretable.	64.55 (24.95)	.610
TF8	I hope that [AI-technology] is clearly explainable.	74.67 (23.56)	.693
CMV Marker Variable Items			
CMV1	Music is important to my life.		
CMV2	Bears are amazing animals.		
CMV3	I find rugby interesting.		
CMV4	When it comes to art, I prefer painting over photography.		

Notes: All items were measured using a 100-point scale (0% totally disagree to 100% totally agree). * Item was dropped during CFA.