Painting A Holistic Picture of Trust in and Adoption of Conversational Agents: A Meta-Analytic Structural Equation Modeling Approach

Eduard Anton
Osnabrück University
eduard.anton@uos.de

Thuy Duong Oesterreich
Osnabrück University
toesterreich@uos.de

Julian Schuir
Osnabrück University
julian.schuir@uos.de

Frank Teuteberg
Osnabrück University
frank.teuteberg@uos.de

Abstract

With their human-like nature, conversational agents (CAs) introduce a social component to human-computer interaction. Numerous studies have previously attempted to integrate this social component by incorporating trust into models such as the technology acceptance model (TAM) to decipher the adoption mechanisms related to CAs. Given the heterogeneity of these previous works, the aim of this paper is to integrate empirical evidence on the role and influence of trust within the nomological network of the TAM. For this purpose, we conduct a meta-analytic structural equation modeling approach based on 45 studies comprising $k = 155$ correlations, and $N = 13,786$ observations. Our findings highlight the multifaceted role of trust as a mediator transmitting the effects of the technology-related perceptions that drive the intention to use CAs. Our results present a comprehensive overview in a thriving research field that can guide both future theory building and the designs of more trustworthy CAs.

1. Introduction

Conversational interactions between humans and computers became an area of interest as early as the 1960s [1]. However, in recent years, technological advancements in computing power and artificial intelligence (AI) algorithms have led to technological maturity and the proliferation of conversational agents (CAs) that are projected to have a market value of nearly $23 billion by 2024 [2]. A CA is software with a text- or voice-based interface that engages with users and exchanges information to serve general purposes or specific domains [3]. Today, CAs are used as digital assistants in a wide range of application areas, such as information retrieval, smart home control and customer service [4]. A CA differs from other information technology (IT) in that, by leveraging natural language processing (NLP), it emulates humanlike behavior and allows users to interact more naturally with technology [5]. This anthropomorphic aspect of CAs alters the common interaction with the software component in sociotechnical systems, raising questions for researchers concerning the conditions under which CAs are adopted or rejected [6, 7, 8].

Given CAs are expected to be useful for many automatable tasks, while improving usability by allowing operations via natural language [9], empirical adoption studies often examine usage behavior by drawing on established technology acceptance theories, such as the technology acceptance model (TAM) [10, 11]. The TAM is a widely used theory in the realm of information systems (IS) that highlights technology-related user perceptions as determinants for adoption intentions in a parsimonious manner [12, 13]. However, Wang and Benbasat [14] have recognized that, in the context of online recommendation agents (an early maturity form of modern CAs), technology-related perceptions are insufficient in capturing usage behavior. This is because users "treat computerized agents as social actors, and form social relationships that involve trust" [14:79]. Therefore, in addition to original TAM constructs such as perceived ease of use and perceived usefulness [13], trusting beliefs (i.e., benevolence, competence, integrity) toward the studied IT artifact are often embedded in empirical research to explain CA adoption [14, 15]. Similar to these studies, we use the term trust to refer to these trusting beliefs as drivers of trusting intentions. In fact, in previous studies, this form of trust (also referred to as trustworthiness [16]) has proven to be an important antecedent of the intention to adopt CAs [17, 18], whereas other studies have revealed no direct effects [6, 19]. This heterogeneity of results can also be observed when examining the importance of other TAM-related constructs that impact CA acceptance. For instance, perceived ease of use affects CA-adoption intentions in the studies of Pal et al. [20] and Pitardi and Marriot [19], but has no significant effect in the studies of Rodríguez Cardona [10], Moussawi et al. [6], and Rese et al. [21]. These inconsistencies among CA-focused studies necessitate further research efforts to achieve an integrated and consistent understanding of the impact of trust within

HICSS

the nomological network of the TAM. Our study addresses this issue by meta-analyzing previous empirical knowledge to advance our understanding of the mechanisms affecting the willingness to interact with CAs. In doing so, we address the call for more trust-related research in the context of CAs [8, 9].

Previous studies synthesizing empirical knowledge within the framework of the TAM neither integrated trust [22, 23] nor had any particular technological focus [12]. Therefore, to the best of our knowledge, this study is the first attempt to provide such an integrated view of CA acceptance research by synthesizing empirical findings across domains and analyzing the variance between studies. Considering the above, we therefore pose the following research question (RQ):

RQ: What is the role and influence of trust on the intention to adopt CAs within the nomological network of the TAM?

We answer this RQ by applying meta-analytic structural equation modeling (MASEM) on the fundament of the Trust-TAM [14, 24]. MASEM combines the strengths of meta-analysis and structural equation modeling (SEM) [25] by enabling us to first conduct a meta-analysis that builds on established empirical findings before fitting the data to the Trust-TAM [14, 24] that we adapted to the CA context. Our goal is not to achieve novelty through the extension of the TAM, nor in the proposal of new constructs, but rather to approach the true effects of trust in conjunction with other TAM-related constructs, thereby strengthening the robustness of previous results and disentangling inconsistencies. These findings can guide the theory building of future studies in a growing body of research and inform design decisions.

The remainder of this paper addresses the theoretical underpinnings of trust in and adoption of CAs (Section 2). We then present the meta-analytical framework of this paper (Section 3) before elaborating on the underlying MASEM methodology (Section 4). The results of the MASEM are subsequently outlined (Section 5). This is followed by a discussion of the research's theoretical and practical implications and its limitations (Section 6). We conclude the study by highlighting the findings of our research (Section 7).

2. Adoption of conversational agents

CAs comprise "both text-based and voice-based automated dialog systems that can interact with a human user via natural language and answer questions on specific topics" [26:2]. From a technical perspective, CAs leverage AI capabilities, including NLP and machine learning, to create a natural conversation flow between humans and machines [27, 28]. As such, CAs partly process users' sensitive and personal information

in an automated manner and, for example, provide recommendations for medical treatments [26]. Since the functionality of CAs is generally substantiated by the storage of large amounts of data, many users are concerned about their privacy when using such CAs [19]. According to a survey conducted by KPMG in 2020, only 28% of US citizens would trust AI-based assistants [29]. In this context, trust refers to "an individual's belief in an agent's competence, benevolence, and integrity" [30:76], which provides the basis for trusting intentions.

Thus far, however, the IS discipline has examined the role of trust in the acceptance of CAs in a fragmented manner, primarily employing unique TAM extensions. For instance, Rodríguez Cardona et al. [10] recently studied the acceptance of CAs within the insurance industry by including privacy concerns and trust into the TAM. Apart from confirming the inherent role of perceived usefulness as the main antecedent regarding the intention to use CAs, the results of their study reveal that trusting beliefs constitute an important precursor of the intention to use CAs, whereas privacy concerns negatively influence perceived trust. In addition, the study suggests that perceived ease of use facilitates the trust-building process. Conversely, an exploratory interview study by Laumer et al. [26] demonstrates that trust in the technology provider of healthcare CAs reduces individual privacy concerns and positively affects the expected utility of a CA. Etemad-Sajadi [18], in turn, found that, apart from the perceived ease of use, trust constitutes an important antecedent in the adoption of virtual agents on websites by positively influencing perceived utility, which in turn forms the main predictor of the intention to use.

In essence, the aforementioned studies indicate that a trustworthy and secure CA design remains necessary for adoption and, in line with the TAM, empirically confirm the inherent role of perceived usefulness [13]. However, other recent studies reveal contradictory results: For instance, Pitardi and Marriot [19] tested a theoretical model in relation to consumers' trust in voice-based CAs such as Amazon's Alexa. Although the authors find that perceived usefulness and ease of use does indeed affect the acceptance of CAs, they demonstrate that trust does not. However, this study indicates that privacy risks do not significantly influence trust in CAs [19]. Additionally, the studies by Moussawi et al. [6] and Rese et al. [21] suggest that a high level of user friendliness does not necessarily promote the adoption of CAs, thereby contradicting the hypothesized relationships of the TAM.

In sum, IS research has so far provided an inconsistent picture of the role and influence of trusting beliefs on the individual adoption intentions of CAs. Considering the increasing diffusion of CAs, we argue

that an integrated comprehension of previous empirical findings can contribute to resolve these inconsistencies and inform future theory building and artifact design. Therefore, our meta-analysis begins by synthesizing the data from extant IS research.

3. Meta-analytical framework

According to Söllner et al. [31], four clusters of trust-based relationships are central to studies in IS research: people-people, people-organizations, organization-organization, and people-technology. Our study is positioned in the latter cluster, which focuses on trust in IT artifacts. This line of literature encompasses many facets and approaches regarding how trust influences the mechanisms of technology adoption, thereby improving our understanding of how to enhance the design of trustworthy artifacts [31]. To this end, trust is often embedded as an external variable in the TAM, with a variety of studies examining its nomological validity for different technologies and contexts [12]. This is not an uncommon procedure, since the TAM, as a key technology acceptance theory, is parsimonious in its original form [13], and empirical research usually extends the theory by including external factors to fit the context under investigation [12].
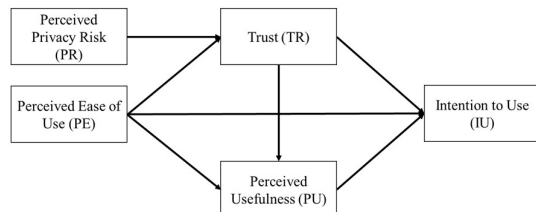

Figure 1. Research model

Our meta-analytical framework, depicted in Figure 1, draws on an earlier attempt to integrate the relational trust perspective with the TAM; Gefen et al. [22] combined both perspectives to study trust in e-vendors in the context of online shopping. This model was later adapted by Wang and Benbasat [16] by considering trusting beliefs toward artifacts in the context of recommendation agents and has been applied to modern CAs [15]. The model adds trust (i.e. trusting beliefs) to the other determinants, such as perceived ease of use and perceived usefulness as antecedents of the intention to use. Perceived usefulness and perceived ease of use thereby constitute the fundamental determinants of use intentions within the TAM [32, 33]. Perceived usefulness refers to "the degree to which a person believes that a particular system would enhance his or her…performance," whereas perceived ease of use indicates "the degree to which a person believes that

using a particular system would be free of effort," [34:320] which in turn affects perceived usefulness [10, 32, 35].

Within the Trust-TAM, trust plays a complementary role as an enabler of perceived usefulness, a mediator of the effects of perceived ease of use [10], and a determinant of the intention to use [17, 18]. Trust is thereby a form of trusting beliefs which are reinforced by perceptions of competence (CAs' competence to effectively achieve the task goal), benevolence (CAs' conveyance that the user's preferences and interests are being pursued), and integrity (CAs' provision of unbiased and honest responses) [16]. For the purpose of the MASEM approach, we considered the reflective latent variable when coding correlations between trust and other constructs (based on the available correlation data in the majority of studies).

To further fit the CA context, we adapt the Trust-TAM [16] and incorporate perceived privacy risk into the meta-analytical framework. This is because the findings concerning modern CAs indicate that risk is a considerable factor affecting trusting beliefs [10, 11], with these beliefs being the precondition of the willingness to take risks (i.e. intention to use) [36]. CAs usually process sensitive and personal data (e.g., health information [11]) that is ultimately processed on servers hosted by specific service providers who seek to gather large amounts of data to render their AI-based services as effective as possible [19]. Thus, many users are concerned about their privacy, which affects their confidence in the benevolence of the technology [19]. As a result, perceived privacy risk when using CAs is generally expected to negatively affect users' trust in CAs.

Our meta-analytic framework serves as a guiding model for the MASEM approach to examine the direct and indirect relationships between the variables.

4. Research methodology

4.1. Literature selection and coding

In February 2021, we searched the Scopus, EBSCOhost, and the AIS Electronic Library databases to comprehensively search for published and unpublished literature. The inclusion of unpublished literature, such as working papers and dissertations, helps mitigate the danger of publication bias, which arises from empirical studies with significant results being more likely to be published than those with non-significant results [37]. The initial database search was conducted by applying the key phrases *((((Conversational OR Virtual OR Voice) AND Agent) OR Chatbot OR "Digital Assistant") AND (experiment*

OR survey OR empirical OR quantitative OR questionnaire) in an article title/abstract/keyword search, yielding in 3,573 hits. After removing 548 duplicates and non-English papers, two researchers independently assessed the relevance of the articles by reading the articles' titles and abstracts. We subsequently examined the full texts to decide whether to include articles.

To be included in the final sample, an article must (1) be exclusively focused on the role of trust in the adoption of CAs, (2) report quantitative findings on the studied relationships that can be extracted in the coding step, and (3) be written in English. As mentioned above, we included both published and unpublished studies for mitigating publication bias and excluded studies that used the same data set (duplicate studies). In addition, we also excluded studies that do not contain original study data (e.g., editorials and research in progress) and studies, that do not report relevant data (e.g., effect size measures). When relevant data was missing (e.g., no correlations were available), we contacted the authors and asked for the missing data. After applying the inclusion and exclusion criteria, we excluded 2,990 articles and identified a total number of 35 relevant articles. We additionally conducted a thorough forward and backward search, which yielded a further 10 relevant articles. Finally, our study sample comprised of 45 articles.[1]

For the coding step, we prepared a coding scheme that enables us to systematically collect the research data prior to the data analysis step. The coding scheme contains relevant research information such as study and reviewer ID, authors, publication year, publication type, sample size, and the effect size measures of our research model's variables of interest, as described in Section 3. To ensure objectivity, consistency, and transparency during the coding step, we defined a set of coding rules which contains detailed guidance on how to code the independent and dependent variables and how to resolve conflicting cases. In addition, we assessed interrater reliability throughout the entire literature selection and coding process. This helped avoid selection biases and enhance the consistency of results, with an overall interrater reliability (Cohen's Kappa) of 0.73, which represents a substantial level of agreement [38].

4.2. Meta-analytic structural equation modeling

The collected data was analyzed in two stages using the MASEM method. In the first stage, the collected correlations of the primary studies are converted into a pooled correlation matrix. In the second stage, this matrix is fitted to the structural research model [25, 39]. This approach is unique in that, although not all of the primary studies report all of the correlations that we consider in our study, we are still able to create an integrated picture from the single pieces of quantitative data [39]. This, in turn, enables us to capture relationships that stem from different models derived from the various theories within our meta-analytic framework [25, 39]. For Stage 1, we adopted a statistical model of random effects. A random effects model stresses a heterogeneous population in which the true effect differs in each primary study. This is because the true effect is often obscured by sampling errors and between-study variances, such as different samples and research designs. The alternative of a fixed-effect model would underline a homogeneous population in which the true effect of the primary studies is only concealed by sampling error [40]. We further followed the meta-analytic paradigm of Hedges and Olkin [41] and did not correct for reliability scores, as this correction procedure (usually conducted in psychometric meta-analyses) does not generally affect the results of MASEM [42].

In light of the random effects model, we consider within- and between-study variance in the calculation of the correlation vector r_i for each study i, so that $r_i = \rho_i + u_i + e_i$, where ρ_i is the correlation matrix, u_i the between-study variance matrix, and e_i the sampling covariance matrix [25]. Leveraging the likelihood estimation method, we then calculated the pooled estimates, yielding the correlation matrix P, its sampling covariance matrix $\hat{v}$, and the between-study heterogeneity T. These pooled estimates form the data basis for Stage 2, in which we fit the data to our proposed structural research model using the weighted least squares (WLS) method, thereby utilizing the asymptotic variances and covariances as the weight matrix [25, 39]. To examine the indirect effects in the structural model, we follow the methodology of Zhao et al. [43].

5. Results

Following the MASEM procedure, we used the combined data of n = 45 studies, k = 155 correlations, and N = 13,786 observations to calculate pooled correlations in Stage 1. The derived meta-analytical correlations of each bivariate relationship are presented in the lower diagonal of Table 1. The resulting relationships range between -.23 (TR-PR) and .63 (PU-IU). All pooled correlations, except some relationships

---

[1] A complete overview of all included individual studies is available upon request.

involving PR (PR-PU, PR-PE, PR-IU), are significantly different from zero given that the 95% confidence intervals (CI) do not enclose zero and the Z scores are larger than the respective critical Z scores of 3.29 [44]. According to Lipsey and Wilson's classification scheme [44], the correlations involving PR can be classified as small, the correlations concerning TR-IU as medium, and the remaining correlations involving either PU or PE as large.

Table 1. MASEM Stage 1 - Results

	PU	PE	PR	TR	IU
PU	-	$\tau^2 =.02$ $I^2=91\%$	$\tau^2 =.07$ $I^2=95\%$	$\tau^2 =.01$ $I^2=76\%$	$\tau^2 =.02$ $I^2=90\%$
PE	.57** n = 33	-	$\tau^2 =.05$ $I^2=93\%$	$\tau^2 =.01$ $I^2=71\%$	$\tau^2 =.02$ $I^2=88\%$
PR	-.14 [n.s.] n = 8	-.05 [n.s.] n = 8	-	$\tau^2 =.12$ $I^2=97\%$	$\tau^2 =.07$ $I^2=95\%$
TR	.61** n = 10	.50** n = 10	-.23* n = 4	-	$\tau^2 =.02$ $I^2=88\%$
IU	.63** n = 33	.51** n = 28	-.13 [n.s.] n = 9	.48** n = 12	-
Lower diagonal: average correlation of bivariate relationship; n = number of studies; significance: **p < 0.001, *p < 0.01, n. s. = not significant. Upper diagonal: τ^2 = between-study variance; I^2 = ratio of between-study variance to overall variance.					

The heterogeneity values indicate a reasonably high dispersion of effects between the studies. This is evident from the Q-test of homogeneity, with a Q-value of 2821.01 (df = 145, p < .01) that exceeds the critical Q-value on the χ^2 distribution. Therefore, the null hypothesis, which states that the dispersion of effects is due to sampling error alone, can be rejected. The Q-test of homogeneity supports our underlying assumption of random effects. The heterogeneity between studies can be expressed by τ^2. However, to obtain a more intelligible value for the proportion of between-study variance to total variance, we added I^2 to the information in the upper diagonal of Table 1, thereby expressing this proportion of the respective bivariate relationship. The values of I^2 vary between 71% and 97%.

The correlation matrix that was meta-analytically derived in Stage 1 is the basis for fitting the data to our research model in Stage 2 by applying the WLS method. We obtained a good model fit with RMSEA and SRMR <.05 [45]. Path analysis (cf. Figure 2) of direct effects revealed only significant effects, except for the relationship TR→IU (β = .13, p > .05). However, TR has a positive effect on PU (β = .44, p < .001) and is, in turn, positively affected by PE (β = .48, p < .001) and negatively by PR (β = -.21, p < .01). The remaining direct relationships also reveal significant effects with PE→PU (β = .35, p < .001), PE→IU (β = .19, p < .001), PU→IU (β = .44, p < .001).

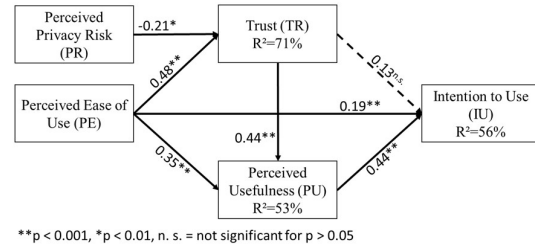**p < 0.001, *p < 0.01, n. s. = not significant for p > 0.05

Figure 2. MASEM Stage 2 - Results

To advance our understanding of the adoption mechanisms in the nomological network, we tested for indirect effects (cf. Table 2). To this end, we follow the procedure of Zhao et al. [43], which starts with establishing the significance of indirect effects. For this purpose, it is particularly advocated to employ the likelihood-based CI method when using MASEM [25, 39]. Significance regarding the indirect effects in the can be derived from the 95% likelihood-based CIs, ranging in the positive value range [25, 39]. The results indicate that PU is fully mediating the effect of TR on IU, as the direct relationship TR→IU is non-significant. In contrast, the effect of PE on PU is partially (complementary) mediated by TR due to the significant direct effect PE→PU. Another partial (complementary) mediation involves the relationship PE→PU→IU [43]. The total effect indicates that PU is the most important determinant of the IU, followed by PE and TR. However, TR plays a key role in strengthening PU in the nomological network.

Table 2. Results of path analysis

Path	DE	Z value	IE	95% LBCI: [LB, UB]	Total effect
PU→IU	.44	4.55	-	-	.44
PE→IU	.19	3.17	.15	[.08, .23]	.34
PE→PU	.35	8.07	.21	[.16, .28]	.56
PE→TR	.48	12.92	-	-	.48
TR→IU	n. s.	.75	.19	[.11, .31]	.19
TR→PU	.44	8.70	-	-	.44
PR→TR	-.21	-2.36	-	-	-.21
DE = direct effect, IE = indirect effect; LB = lower bound of the 95% likelihood-based CI; UB = upper bound of the 95% likelihood-based CI; Total effect = direct effect + indirect effect; n. s. = not significant.					

6. Discussion

By synthesizing the results of 45 CA studies, including 155 correlations and 13,786 empirical observations, through a MASEM approach, we contribute to enhancing the understanding of how trusting beliefs affect the adoption mechanisms of CAs.

We adapted the Trust-TAM [14, 24] as a meta-analytic framework that merges the trust perspective with the TAM's technological perceptions of CAs. We thereby tested and highlighted the interrelations explaining CA adoption. Our findings present a more consistent and holistic understanding of trust in affecting the intention to adopt CAs in an emerging and dynamic research field with heterogeneous findings, thereby contributing to guiding future theory building and helping practitioners design trustworthy CAs. We address the RQ by summarizing the main findings (MF) and the corresponding implications for research (IR) and practice (IP) in Table 3. Furthermore, we elaborate on these findings in light of the theoretical background, the limitations of our study, and the possible avenues for future research.

Table 3. Main findings and implications

Research question: What is the role and influence of trust on the intention to adopt CAs within the nomological network of the TAM?
MF1: Trust has no direct effect on the intention to use; instead, the relationship between trust and the intention to use is fully mediated by perceived usefulness. In particular, trust positively affects the perceived usefulness, which in turn has a strong impact on the intention to use.   • IR1.1: More research is necessary to further comprehend the role of trust in affecting the intention to use CAs in different contexts and domains.   • IR1.2: Future meta-analyses should address the moderating effects of the application context on this relationship.   • IR/IP1.3: Scholars and practitioners should focus further on promoting the system's usefulness when designing and introducing CAs.
MF2: Perceived ease of use positively affects the intention to use CAs and has a strong impact on trust.   • IP2.1: System designers should integrate functional and design elements into the user interface that improve the usability of the technology when designing CAs.   • IR2.2: Future research efforts should focus on exploring the role of human-like aspects of CAs in affecting trust and adoption intentions.   • IR2.3: Future meta-analyses should integrate the humanness level of CAs and usage contexts into their analyses as further moderator variables.
MF3: Trust mediates the strong direct effect of the perceived ease of use on perceived usefulness in a complementary manner. Thus, the perceived ease
of use is both an enabler of trust and an important determinant of perceived usefulness.   • IR/IP3: System developers should integrate design principles that enable a goal-oriented, unambiguous conversational flow through improved confirmation and error-handling strategies [28].
MF4: Perceived privacy risk is inhibiting the trustworthiness of CAs.   • IP4.1: Policy makers should direct more focus on legal issues to protect users' data and privacy, thereby finding a balance between an appropriate level of privacy protection and a legal environment that provides space for leveraging the technology [46].   • IR4.2: Future meta-analysis should incorporate the moderating influences of cultural diversity when assessing the impact of privacy-risk effects on usage intentions.   • IR4.3: More research is required to compare trust levels in different trust targets in the context of privacy concerns (e.g., trust in the technology versus the provider).

6.1. Implications for research and practice

MF1: Our results reveal that trust only indirectly affects the usage intentions of CAs via perceived usefulness. When users place trusting beliefs in a technology such as a CA, they perceive this technology as being useful. Users' perceptions that the technology is helpful for completing their task, in turn, strongly affect their intention to use, as is revealed by the TAM [13]. The meta-analytical results of our study support the notion in the literature that several usage contexts do not demand trust in technology. For example, Pitardi and Marriot [19] argue that trust in CAs is unnecessary because they are primarily used for simple tasks (e.g., for smart home applications such as turning lights on and off). However, in higher-risk settings, such as insurance [10], results from primary studies suggest small but direct effects, as the stakes are higher for the user. Trust has been emphasized as being highly context sensitive [36] and anecdotal evidence has indicated that user behavior varies depending on whether IT is applied in a utilitarian or hedonic context [47]. Three major IRs arise from MF1. First, more research is required to effectively comprehend the role of trust in affecting the intention to use CAs in different contexts and domains. Our literature search has revealed that the number of primary studies is currently insufficient for an extended meta-analytic context analysis per domain; for example, in health, insurance, and home contexts *(IR1)*. Second, given the heterogeneity among individual CA studies,

future meta-analyses should address the moderating effects of the application context to gain deeper insights into usage contexts in which trust is somewhat salient depending on the availability of sufficient primary CA studies *(IR2)*. Third, scholars and practitioners should focus on promoting and emphasizing the system's usefulness when designing and introducing CAs *(IR/IP3)*. As implied by MF1, the usefulness of CAs plays a crucial role in affecting the adoption intention. Thus, in addition to the highlighted role of trust in conceptualizing the benevolence, competence, and integrity level of CAs [48, 49], the usefulness of the technology must be addressed to enhance the adoption intention in the IT field. In this context, Rodríguez Cardona et al. [10] recommend a "trade-off between practical use and trust," since the willingness to interact with a CA corresponds directly to the importance for a user of the use case. Thus, CA developers should direct their attention to improving the system's functionality by validating CA designs, assessing the effectiveness of task fulfillment, and conducting technology acceptance questionnaires to improve the user experience [50], thereby increasing the transfer of trustworthiness.

MF2: Contrary to the studies indicating that perceived ease of use has no effect on CA adoption [6, 10, 21], our integrated results provide evidence of a direct relationship between these variables. Moreover, we observed a strong direct impact of the perceived ease of use on trust, indicating that users tend to place trust in CAs that are perceived as easy to use. Perceived ease of use describes the perception of the effort required to use CAs [13]. Various studies have demonstrated that specific design elements can reduce effort in interacting with CAs, and thus increase usage satisfaction. For example, elements such as auto suggestion buttons can be integrated into text-based interfaces to reduce typing effort [51] In the case of voice assistants, predefined response options can guide the user through dialog [5]. Thus, MF2 implies that system designers should integrate functional and design elements into the user interface that help to improve the usability of the technology when designing CAs (IP.2.1).

Moreover, technological humanness is an issue that deserves more attention in CA research when studying the relationship between perceived ease of use, trust, and CA usage intentions. Technological humanness has recently emerged as a potent field of research that focuses on human-like design elements for CAs that render the human-computer interaction more natural and thus more effortless [28, 52]. A comprehensive meta-analysis conducted by Blut et al. [52] suggests that perceived ease of use acts as an operating mechanism of anthropomorphic design elements and mediates their effects on the intention to use CAs. Therefore, our meta-analytic framework nurtures the integration of further

perspectives, such as the uncanny valley theory or the social presence theory, to explore the role of human-like aspects of CAs in building trusting beliefs and enhancing usage intentions in different contexts. This would contribute to extending the relational trust perspective and the usability aspects of the TAM (IR2.2).

Prior studies have provided evidence for the major role of social presence as an important determinant of trust when interacting with CAs, such as Alexa [19, 53]. However, anecdotal evidence has also suggested that the impact of human-like characteristics depends on specific contexts [50]. In particular, the human-like characteristics of a technology are considered helpful in building trust when the system is intended to be used as a surrogate for a human, whereas they are not helpful when the system is intended to replace a machine [50]. Future meta-analyses should elaborate on this aspect by integrating the humanness level of CAs and usage contexts as further moderator variables (IR2.3). Deeper insights into usage contexts in which a particular level of humanness is required to generate trust would help system designers design trustworthy CAs and to increase usage intentions. For example, there is an interesting question to be answered regarding which level of humanness is necessary in which usage contexts.

MF3: Apart from the impact of ease of use on trust and use intention, results indicate that the effects of an effortless engagement with CAs on adoption intentions are indirectly transmitted through the perceived usefulness, and its effects on perceived usefulness are mediated through trust. This suggests that "even when the design of a [CA] is simple, clear, understandable and easy to learn, if the operational outcomes and expected benefits do not match the user's cognitive trust expectations, the [CA] will not contribute to optimize the customer experience" [10]. Thus, perceived ease of use acts as a strong enabler of trustworthiness and perceived usefulness. However, perceived usefulness is ultimately the critical determinant that drives adoption. Hence, aside from the implications derived from MF2 (IP2.1/IR2.2/IR2.3) and, in line with suggestions from recent CA studies [28], we recommend developers adhere to effectiveness-enhancing design principles that ensure a goal-oriented, unambiguous conversational flow that takes into account confirmation and error-handling strategies (IR/IP3).

MF4: While privacy risk is not a part of the original Trust-TAM [14, 24], we included the construct in our meta-analytical framework, as prior studies frequently highlight privacy concerns that arise from the use of CAs and empirically confirm the negative effect of privacy risk on trust [10, 11]. Our meta-analytical analysis confirms the inhibiting role on trust in CAs.

Although data protection laws have been tightened in regions such as Europe following the GDPR, users are still reluctant to entrust their data to CAs [19]. Legislatures must continue seeking ways to protect users' data and privacy regardless of where the servers are located (IP4.1). However, an excessive level of data protection restricts the capabilities of AI, necessitating a balance between privacy protection (e.g., pseudonymization, cyber security measures), legal requirements, and the use of data in leveraging the technology [46].

In addition to legislative differences, privacy risks need to be assessed against the background of cultural diversity. Cultures, such as those in the Asian region, tend to exhibit more optimistic attitudes toward AI compared to cultures, such as those in the European region, which tend to be more averse to new technologies [54]. Future meta-analysis should incorporate such moderating influences when assessing the impact of privacy-risk effects on usage intentions (IR4.2). A deeper understanding of the circumstances under which cultural diversity accounts for differing risk perceptions would help system developers and managers design and establish country-specific levels of data protection.

Furthermore, Pitardi and Marriott [19] argue that the level of privacy concern may depend on whether the user perceives the CA as a standalone entity or as the provider behind the CA (e.g., Amazon behind Alexa). This perspective stems from conceptual differences between trust in the technology and trust in the provider [55]. In the latter, the trustee is no longer the technological artifact, but the vendor managing the technology and the data. Thus, the trust level can vary significantly between these perspectives and can cause between-study variances [19]. Söllner et al. [56] argue that multiple targets of trust relationships matter to understanding adoption behavior regarding a technology. Therefore, we propose that future studies further explore institution-based trust in the context of privacy concerns to compare trust levels in different trust targets (IR4.3).

6.2. Limitations

The applied methodological approach is not free of limitations. As with any meta-analytic approach, there is the issue of the "apples and oranges" problem that challenges the validity of results. This is because the approach incorporates different research results, measures and variables with varying research designs and populations [57]. Since the purpose of meta-analysis is to synthesize research findings, the key is to find the appropriate scope of synthesis. We followed the recommended guidelines and set a specific research

focus within the CA-adoption literature and identified the appropriate samples through a systematic literature review [57]. In doing so, we ensured that we are "combining apples and oranges to understand something about fruit [which] may make more sense than combining fruit and humans to understand something about organic matter" [58:20].

Another limitation concerns heterogeneity, which suggests moderator effects. We did not perform a moderator analysis primarily because of the unavailability of sufficient primary studies. Moderator analysis in MASEM approaches requires a larger number of studies because grouping (e.g., contextual moderator analysis regarding utilitarian and hedonistic application domains) divides the sample size into smaller subgroups for MASEM analysis [25]. Future research should, first, continue to conduct primary research in different contexts and, second, examine meta-analytic moderator influences with an expanded sample.

Finally, we would stress that the studied form of trust is trusting beliefs, which is distinct from trusting intentions and trusting behavior [55]. This studied form of trust is often used synonymously with trustworthiness [16]. These trust distinctions and terminologies are often vague and not clearly defined in the examined IS literature. Future research should pay attention to the distinction and the terminology of different trust forms.

7. Conclusion

By placing the focus on CAs, we revisited the Trust-TAM [14, 24], and undertook an integrated examination of the adoption mechanisms of modern voice- and text-based agents. Using the MASEM approach, we investigated how trust simultaneously interacts with perceived usefulness, perceived ease of use, and perceived privacy to influence intentions to use CAs. Our findings demonstrated that trust does not directly influence usage intentions, but instead unfolds fully via perceived usefulness. Nevertheless, trust occupies a multifaceted role as a mediator that transfers the effects of perceived ease of use to perceived usefulness. Trust also functions as an interceptor of the inhibiting effects of perceived privacy risks. In addition, our results indicated that the most important criteria concerning adoption intentions are the perceived usefulness of CAs in automating tasks or achieving a specific goal, and the perceived ease of use through natural language.

With these insights into a variety of different CAs, use cases, and populations, we were able to contribute to a deeper understanding of the principal trust-building mechanism within the Trust-TAM and aid future theory building. This enhanced understanding helps scholars to

identify future research opportunities and practitioners to direct their attention toward issues that are useful in designing trustworthy CAs. In essence, we were able to indicate that building trustworthy CAs is amplified by their ability to improve users' effectiveness and productivity. With this meta-analysis study, we were able to offer a holistic and consistent understanding of trust in and adoption of conversational agents based on prior CA studies from an emerging and dynamic research field.

8. References

[1] Weizenbaum, J., "ELIZA—a computer program for the study of natural language communication between man and machine", *Communications of the ACM 9*(1), 1966, pp. 36–45.

[2] Overend, R., S. Stefanoff, P. Jackson, S. Rayment, S. Cooper, and D. Taylor, *Conversational AI*, 2019.

[3] Diederich, S., A.B. Brendel, and L.M. Kolbe, "Towards a Taxonomy of Platforms for Conversational Agent Design", *14th International Conference on Wirtschaftsinformatik*, (2019).

[4] Laumer, S., A. Racheva, F. Gubler, and C. Maier, "Use Cases for Conversational Agents : An Interview-based Study", *AMCIS*, (2019).

[5] Diederich, S., A.B. Brendel, S. Lichtenberg, and L.M. Kolbe, "Design for fast request fulfillment or natural interaction? Insights from an experiment with a conversational agent", *Twenty-Seventh European Conference on Information Systems (ECIS2019)*, (2019).

[6] Moussawi, S., M. Koufaris, and R. Benbunan-Fich, "How perceptions of intelligence and anthropomorphism affect adoption of personal intelligent agents", *Electronic Markets*, 2020, pp. 1–22.

[7] Laumer, S., C. Maier, and F.T. Gubler, "CHATBOT ACCEPTANCE IN HEALTHCARE: EXPLAINING USER ADOPTION OF CONVERSATIONAL AGENTS FOR DISEASE DIAGNOSIS", *Proceedings of the 27th European Conference on Information Systems (ECIS)*, (2019).

[8] Zierau, N., C. Engel, M. Söllner, and J.M. Leimeister, "Trust in smart personal assistants: A systematic literature review and development of a research agenda", *15th International Conference on Wirtschaftsinformatik*, (2020).

[9] Diederich, S., A.B. Brendel, S. Morana, and L. Kolbe, "On the Design of and Interaction with Conversational Agents: An Organizing and Assessing Review of Human-Computer Interaction Research", *Journal of the Association for Information Systems forthcomin*, 2022.

[10] Rodríguez Cardona, D., A. Janssen, N. Guhr, M.H. Breitner, and J. Milde, "A Matter of Trust? Examination of Chatbot Usage in Insurance Business", *Proceedings of the 54th Hawaii International Conference on System Sciences*, (2021).

[11] Prakash, A.V., and S. Das, "Would you Trust a Bot for Healthcare Advice? An Empirical Investigation", *Twenty-Fourth Pacific Asia Conference on Information Systems*, (2020).

[12] Wu, K., Y. Zhao, Q. Zhu, X. Tan, and H. Zheng, "A meta-analysis of the impact of trust on technology acceptance model: Investigation of moderating influence of subject and context type", *International Journal of Information Management 31*(6), 2011, pp. 572–581.

[13] Davis, F.D., "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology", *MIS Quarterly 13*(3), 1989, pp. 319.

[14] Benbasat, I., and W. Wang, "Trust In and Adoption of Online Recommendation Agents", *Journal of the Association for Information Systems 6*(3), 2005, pp. 72–101.

[15] Müller, L., J. Mattke, C. Maier, T. Weitzel, and H. Graser, "Chatbot acceptance: A latent profile analysis on individuals' trust in conversational agents", *SIGMIS-CPR 2019 - Proceedings of the 2019 Computers and People Research Conference*, 2019, pp. 35–42.

[16] Colquitt, J.A., B.A. Scott, and J.A. LePine, "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), 2007, pp. 909–927.

[17] Kasilingam, D.L., and A. Soundararaj, "Are Chatbots Going to Replace Mobile Shopping Applications? Understanding the Attitude and Intention to Use Chatbots for Shopping Using Smartphones", *SSRN Electronic Journal*, 2020.

[18] Etemad-Sajadi, R., "The influence of a virtual agent on web-users' desire to visit the company", *International Journal of Quality & Reliability Management 31*(4), 2014, pp. 419–434.

[19] Pitardi, V., and H.R. Marriott, "Alexa, she's not human but… Unveiling the drivers of consumers' trust in voice-based artificial intelligence", *Psychology & Marketing 38*(4), 2021, pp. 626–642.

[20] Pal, D., C. Arpnikanondt, S. Funilkul, and W. Chutimaskul, "The Adoption Analysis of Voice-Based Smart IoT Products", *IEEE Internet of Things Journal 7*(11), 2020, pp. 10852–10867.

[21] Rese, A., L. Ganster, and D. Baier, "Chatbots in retailers' customer communication: How to measure their acceptance?", *Journal of Retailing and Consumer Services 56*(June), 2020, pp. 102176.

[22] Wu, J., and A. Lederer, "A Meta-Analysis of the Role of Environment-Based Voluntariness in Information Technology Acceptance", *MIS Quaterly 33*(2), 2009, pp. 419–432.

[23] Turner, M., B. Kitchenham, P. Brereton, S. Charters, and D. Budgen, "Does the technology acceptance model predict actual use? A systematic literature review", *Information and Software Technology 52*(5), 2010, pp. 463–479.

[24] Gefen, Karahanna, and Straub, "Trust and TAM in Online Shopping: An Integrated Model", *MIS Quarterly 27*(1), 2003, pp. 51.

[25] Cheung, M.W.L., *Meta-Analysis*, John Wiley & Sons, Ltd, Chichester, UK, 2015.

[26] Laumer, S., C. Maier, and G. Fabian, "Chatbot acceptance in healthcare: Explaining user adoption of conversational agents for disease diagnosis", *27th European Conference*

on Information Systems - Information Systems for a Sharing Society, ECIS 2019, (2019), 1–18.

[27] Feine, J., U. Gnewuch, S. Morana, and A. Maedche, "A Taxonomy of Social Cues for Conversational Agents", *International Journal of Human Computer Studies*, 2019.

[28] Gnewuch, U., S. Morana, and A. Maedche, "Towards Designing Cooperative and Social Conversational Agents for Customer Service", *Thirty Eighth International Conference on Information Systems*, (2017).

[29] Gillespie, N., S. Lockey, and C. Curtis, *Trust in artificial Intelligence: a five country study*, 2021.

[30] Wang, W., I. Benbasat, and W. Wang, "Trust and TAM for Online Recommendation Agents Trust and TAM for Online Recommendation Agents", 2004.

[31] Söllner, M., D. Gefen, J.M. Leimeister, and P.A. Pavlou, "Trust: An MIS Quarterly research curation", *MIS Quarterly*(October), 2016, pp. 1–9.

[32] Patil, K., and M.S. Kulkarni, "Artificial Intelligence in Financial Services: Customer Chatbot Advisor Adoption", *International Journal of Innovative Technology and Exploring Engineering 9*(1), 2019, pp. 4296–4303.

[33] Almahri, F.A.J., D. Bell, and M. Merhi, "Understanding Student Acceptance and Use of Chatbots in the United Kingdom Universities: A Structural Equation Modelling Approach", *2020 6th International Conference on Information Management (ICIM)*, IEEE (2020), 284–288.

[34] Davis, F.D., "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology", *MIS Quarterly 13*(3), 1989, pp. 319.

[35] Pal, D., C. Arpnikanondt, S. Funilkul, and M.A. Razzaque, "Analyzing the adoption and diffusion of voice-enabled smart-home systems: empirical evidence from Thailand", *Universal Access in the Information Society*(November), 2020.

[36] Mayer, R.C., J.H. Davis, and F.D. Schoorman, "An Integrative Model Of Organizational Trust", *Academy of Management Review 20*(3), 1995, pp. 709–734.

[37] Rothstein, H.R., A.J. Sutton, and M. Borenstein, *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments*, Wiley, 2005.

[38] Landis, J.R., and G.G. Koch, "The Measurement of Observer Agreement for Categorical Data", *Biometrics 33*(1), 1977, pp. 159.

[39] Jak, S., *Meta-Analytic Structural Equation Modelling*, Springer International Publishing, Cham, 2015.

[40] Borenstein, M., L. V. Hedges, J.P.T. Higgins, and H.R. Rothstein, *Introduction to Meta-Analysis*, John Wiley & Sons, Ltd, Chichester, UK, 2009.

[41] Hedges, L. V., and I. Olkin, *Statistical Methods for Meta-Analysis*, Academic Press, Inc., 1985.

[42] Michel, J.S., C. Viswesvaran, and J. Thomas, "Conclusions from meta-analytic structural equation models generally do not change due to corrections for study artifacts", (December 2017), 2011.

[43] Zhao, X., J.G. Lynch, and Q. Chen, "Reconsidering Baron and Kenny: Myths and Truths about Mediation Analysis", *Journal of Consumer Research 37*(2), 2010, pp. 197–206.

[44] Lipsey, M.W., and D.B. Wilson, "Practical meta-analysis", *Applied Social Research Methods Series*, 2001.

[45] Hu, L., and P.M. Bentler, "Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives", *Structural Equation Modeling: A Multidisciplinary Journal 6*(1), 1999, pp. 1–55.

[46] He, J., S.L. Baxter, J. Xu, J. Xu, X. Zhou, and K. Zhang, "The practical implementation of artificial intelligence technologies in medicine", *Nature Medicine 25*(1), 2019, pp. 30–36.

[47] Gerow, J.E., R. Ayyagari, J.B. Thatcher, and P.L. Roth, "Can we have fun @ work? the role of intrinsic motivation for utilitarian systems", *European Journal of Information Systems 22*(3), 2013, pp. 360–380.

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

[49] Chae, S.W., K.C. Lee, and Y.W. Seo, "Exploring the Effect of Avatar Trust on Learners' Perceived Participation Intentions in an e-Learning Environment", *International Journal of Human–Computer Interaction 32*(5), 2016, pp. 373–393.

[50] Diederich, S., A.B. Brendel, and L.M. Kolbe, "Designing Anthropomorphic Enterprise Conversational Agents", *Business & Information Systems Engineering 62*(3), 2020, pp. 193–209.

[51] Jain, M., P. Kumar, R. Kota, and S.N. Patel, "Evaluating and Informing the Design of Chatbots", *Proceedings of the 2018 Designing Interactive Systems Conference*, ACM (2018), 895–906.

[52] Blut, M., C. Wang, N. V. Wünderlich, and C. Brock, "Understanding anthropomorphism in service provision: a meta-analysis of physical robots, chatbots, and other AI", *Journal of the Academy of Marketing Science*, 2021.

[53] Yen, C., and M.-C. Chiang, "Trust me, if you can: a study on the factors that influence consumers' purchase intention triggered by chatbots based on brain image evidence and self-reported assessments", *Behaviour & Information Technology*, 2020, pp. 1–18.

[54] Duranton, S., J. Erlebach, and M. Pauly, *Mind the (AI) Gap: Leadership Makes the Difference*, BCG GAMMA, 2018.

[55] McKnight, D.H., V. Choudhury, and C. Kacmar, "Developing and Validating Trust Measures for e-Commerce: An Integrative Typology", *Information Systems Research 13*(3), 2002, pp. 334–359.

[56] Söllner, M., A. Hoffmann, and J.M. Leimeister, "Why different trust relationships matter for information systems users", *European Journal of Information Systems 25*(3), 2016, pp. 274–287.

[57] Sharpe, D., "Of apples and oranges, file drawers and garbage: Why validity issues in meta-analysis will not go away", *Clinical Psychology Review 17*(8), 1997, pp. 881–901.

[58] Hall, J.A., R. Rosenthal, I. Tickle-Degnen, and F. Mosteller, "Hypotheses and problems in research synthesis", In H. Cooper and L. V. Hedges, eds., *The handbook of research synthesis*. Russel Sage, New York, 1994, 17–28.