What Are the Factors That Drive AI Acceptance: A Meta-Analysis Approach

Aslı Gül Kurt[1,2](✉), Alexander John Karran[2], Ruxandra Monica Luca[1,2], and Sylvain Sénécal[1,2]

[1] HEC Montreal, Montreal, QC H3T 2A7, Canada
[2] Tech3lab, HEC Montreal, Montreal, QC H3T 1T7, Canada
{asli-gul.kurt,alexander-john.karran,ruxandra-monica.luca, sylvain.senecal}@hec.ca

Abstract. Antecedents of technology acceptance (TA) are known to be positively associated with measures such as usage intention, behavioral intention, attitude, and satisfaction. Although technology acceptance is investigated widely in prior research, it is not currently clear which variables or factors drive technology acceptance and under different service contexts or conditions. To examine the strength these effects in the artificial intelligence literature, we adopt a meta-analysis approach. We have scoped the literature on artificial intelligence, acceptance measures, and factors affecting acceptance in extant literature. We narrowed our search to business context to find AI-based tools that users, consumers, and customers interact with transactionally, such as chatbots. Findings show AI-based technology factors affect acceptance differently in various service industry contexts as preliminary results. These results have critical implications for researchers and practitioners studying which type of AI-based technology strengthen consumers use in different service contexts. These preliminary findings will be extended to look at interactive relationships of factors affecting acceptance in different contexts.

Keywords: Technology acceptance · Artificial intelligence · Meta-analysis · AI factors

1 Introduction

Artificial intelligence (AI) technologies are increasingly utilized across platforms and in different service contexts. However, understanding the degree to which they are accepted, and under which circumstances, requires further investigation. Traditionally, within the information systems research, technology acceptance (TA) has been extensively studied through different models. These models include a plethora of variables that precede individual responses, such as usage intention, behavioral intention, attitude, and satisfaction towards the technology. Two of the most frequently utilized models are the Unified Theory of Acceptance and Use of Technology (UTAUT) [13], and the Technology Acceptance Model (TAM) [2]. Moreover, users' TA level is also affected by the presence of anthropomorphism, perceived trust, risk, privacy, enjoyment, and satisfaction.

J. Y. C. Chen et al. (Eds.): HCII 2022, LNCS 13518, pp. 329–337, 2022.
https://doi.org/10.1007/978-3-031-21707-4_24

First of these set of factors is anthropomorphism, which is the indication of having the human-like characteristics present in technology, such as a human appearance, emotions, or intentions. Previous research has shown that AI technologies that primarily interact with humans like chatbots [11] and recommendation agents [9] have better adoption and intention to use rates when the agent has anthropomorphic features. The second factor, trust, and its antecedents in artificial intelligence technologies are highly investigated [8]. These investigations include performance and attribute-related factor such as system competencies or personality and communication. Aside from these performance and attribute-based factors of AI agents or systems, enjoyment and satisfaction are also factors that have been shown to increase an individual's intentions to adopt or use the technology [5,15]. Other factors related to the features of an artificial intelligence system are the risk and privacy levels perceived by humans. A growing stream of research has shown that these features that can negatively affect individuals' trust in the system lead to lower continued use [4,7]. Apart from these factors, AI systems have been documented to be affected by other attributes as explainability and emotional conversationality, which in turn can reduce an individual's cognitive load [6], help build better interactions with the technology, and lead to a more natural type of communication [1].

Although these attributes, which include the conversational, emotional, and explainable factors of AI, have been studied across many disciplines. It is not yet clear which one of these factors leads to the strongest levels of acceptance. Thus, the goal of this research is to investigate the factors driving TA in AI systems, through a meta-analysis approach.

2 Methods

2.1 Finding Sample of Articles

To complete the literature search, we follow the guidelines from past meta-analyses done on both AI and acceptance models [8]. We conducted our first search on the Web of Science platform, which covers the majority of the publishers and extends to multiple fields of research. We developed an initial search term that includes various chatbot systems and acceptance terms that will correlate to our interest in this meta-analysis. We have started with key terms such as "chatbot", "UTAUT", "TAM", and "usage intention" or "adoption intention" and revisited our search term in each iteration of title and abstract screening.

Our keyword search term(s) were based on acceptance and AI keywords, which comprised AI factors within or without technology acceptance models, adoption or usage intention constructs, and AI keywords used in business contexts, respectively. We also developed some inclusion criteria regarding the publication year and language of the articles. AI research has transformed quite in recent years; consequently, the decision was made to include the research done between 2021–2022 in English. Furthermore, for specific research fields of interest, the search terms were narrowed include only "Information systems", "Business", and "Management" fields. Inclusion criteria included empirical research that has been done within our scope of research questions.

For our following round of literature search, we have selected ScienceDirect as a platform since it included research from consumer research and the social sciences as business or management compared to other platforms. For this platform, we have narrowed our search term(s) to keywords such as "chatbot", "adoption intention", "usage intention", and included more general terms such as "usage" or "adoption". These additional inclusion criteria were inserted to eliminate any additional noise that may emerge from searching for other AI terms. Per our previous search using the Web of Science platform, this search also included research articles published in English within the last decade. We did not include the specification for business-related fields for this search to find articles that fit our research questions from other industries such as health, tourism, etc. From both searches, there were 1633 and 644 returns from Web of Science and ScienceDirect respectively, resulting in 2277 articles to screen. Due to time limitations at hand, we started the title and abstract screening process with extraction of the data. After the screening process for title and abstracts of the results returned from these search terms, we have extracted 18 articles for our preliminary search, following the PRISMA guidelines (Fig. 1).

2.2 Selection of Articles

For the next stage of our review, we have created an inclusion criteria list. After screening for every criterion within our search returns, 18 articles and total of 87 effect sizes were identified and extracted. Although there were more records returned after our search, only 18 were selected. The code-book and article information coded are available upon contact with corresponding author. The reason we selected these articles was to be able to show a preliminary set of results. Articles were coded following the Coding Scheme in Appendix A (Fig. 2). The criteria followed for this process to include articles may also be found in Appendix B.

2.3 Analysis Protocol of the Effect Sizes

The pre-processing stage consisted of converting the results from the statistical tests and regression models to effect sizes. Due to the fact that all 18 articles included in the analysis did not report any effect sizes, we converted them manually. The majority of the articles reported t-values of the relationships between AI acceptance characteristics and their precedent usage, adoption intention, attitude, or satisfaction. Only nine articles reported p-values for their tests. While the one remaining study did not report t-values, it included β and standard errors from the regression model, which can be converted to t-values. The transformations for all the values are presented below:

$$d = \frac{(2 * r)}{\sqrt{(1 - r^2)}}, \tag{1}$$

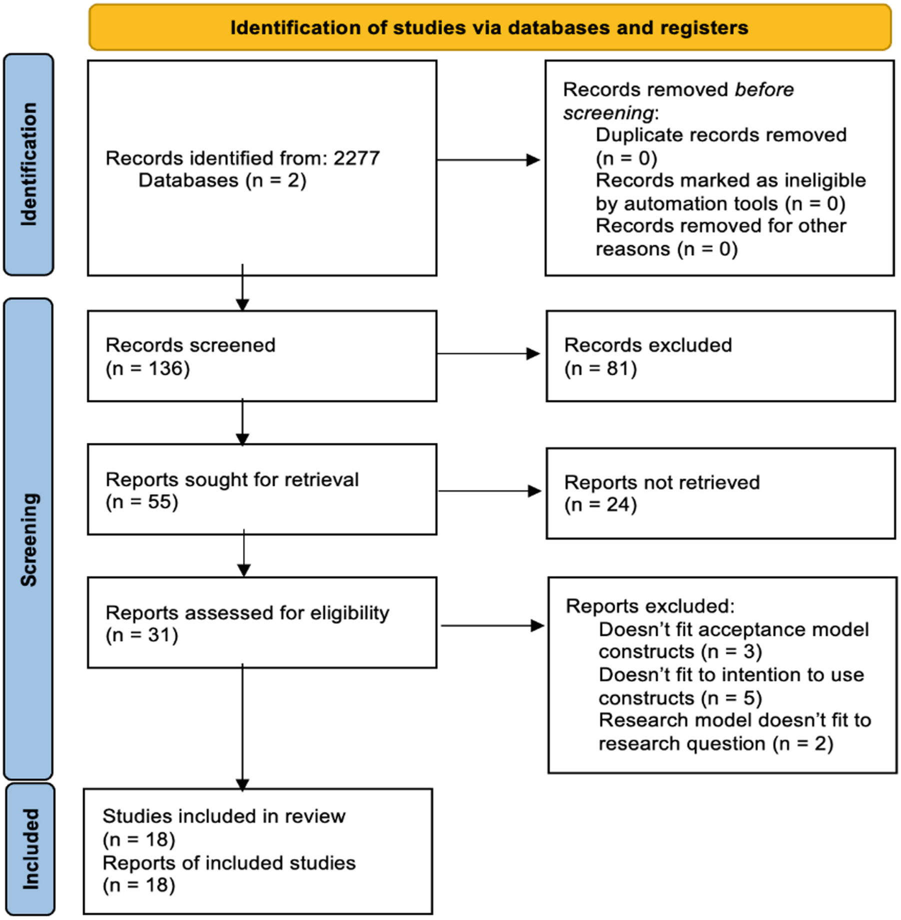

Fig. 1. PRISMA flowchart for identifying studies via databases for meta-analysis.

$$d = \frac{t}{\sqrt{(t^2 + n - df)}}, \tag{2}$$

$$t = \frac{\beta}{SE}, \tag{3}$$

where d and r are the effect size Cohen's d and r, t is the t-value, n sample size, and β and SE are the coefficient and the standard error of the structural equation model results.

The analysis for the relationship between characteristics of acceptance and the precedent factors (i.e. usage, adoption, attitude, and satisfaction) is the main relationship of interest. Others include the role of conversational and emotional attributes, the type of industry, and finally the personal versus professional uses of AI on effect size strength. In order to have a reasonable amount of variation in the analysis, there were at least five effect sizes for each moderator. Furthermore, we only extracted the total or direct effects of interest from articles that included structural equation modelling and were relevant to our primary research question. Finally, since this is a presentation of preliminary results, p-values less than 0.1 will be taken into consideration as marginally significant. To test the effects of the interaction of interest with moderators we will use R package metaphor [14], which will help execute these tests with a random-effect multi-level meta-regression model.

3 Results

3.1 Homogeneity Analysis and Publication Bias

As is the standard procedure for a meta-analysis article, we completed several analyses. Firstly, we performed a homogeneity test of the sample, given that we are investigating the presence of any moderators or other variables that might produce the variability within the effect sizes. The homogeneity test revealed there were indeed moderating factors associated with the effect of factors on acceptance ($Q(86) = 1565.81$, $p < .0001$) within our sample. Finally, Egger's test [3], trim and fill [12], and Rosenthal [10] methods revealed no asymmetry within our sample, suggesting no publication bias.

3.2 Factors Driving Acceptance

A meta-regression model including all our moderators showed that the AI factor that driving a strong effect on acceptance is conversational ($r = 1.79$, $p = 0.0177$, CI = [0.188, 1.979]) and explainable ($r = 0.81$, $p = 0.0726$, CI = [-0.045, 1.033]) attributes of artificial intelligence, compared to the traditional measures of acceptance.

We further investigated whether explainable, conversational, and emotional attributes are more useful in some industries than others. Based upon our findings, having explainable features in AI interacts positively with acceptance in both e-commerce ($p < .0001$, CI = [-1.1023, -0.5764]) and tourism ($p < .0001$, CI = [0.6395, 1.235]). Although we did not find any significant interaction between conversational attributes and industry types, having this attribute in an artificial intelligence system produces larger effect sizes than not having it.

Similarly, an interaction effect between the emotionality of the agent and acceptance within each industry was not significant except e-commerce, which produced stronger effects when there were emotional cues made by the AI ($p = 0.0172$, CI = [0.0921, 0.9492]). When it comes to personal and professional use of these AI agents in service context, our initial results also showed interesting but non-significant results. Based on our results, in professional settings, all three attributes (conversational, emotional, and explainable) positively affected acceptance compared to when these attributes were absent. Moreover, emotionality was the only factor that increased acceptance in personal use compared to the other factors.

3.3 Robustness Check

To be able to do a robustness check, we have disattenuated the reliability scores that are used in each study to measure independent and dependent variables. These reliability scores are collected from these variables ranged from 0.63 to 0.98, and 0.76 to 0.98, respectively. After the calculation of disattenuated r and variance of disattenuated rs, these values are used to do an additional robustness check. This model also showed us that there may be heterogeneity in our sample ($Q(86) = 3470.88$, $p < .0001$).

4 Discussion

Through our initial analysis of the literature, we have identified which factors play an important role in the TA literature. We found that different industries that utilize AI technologies have different needs when it comes to certain factors of AI. Furthermore, emotional, conversational, and explainable factors do not affect TA in AI artifacts unanimously. In other words, employing different factors to different sectors are crucial to individuals' acceptance rates.

Surprisingly, when the effect sizes produced within these articles were compared, factors such as satisfaction and privacy had a stronger effect on TA than the traditional acceptance models. After investigating these relationships across industries, emotional aspects of AI were found to negatively affect acceptance in sectors that could increase vulnerability or include more vulnerable populations. The health sector is one such example. On the other hand, the presence of conversational factors of AI do not appear to affect acceptance in other sectors such as e-commerce and tourism.

4.1 Limitations

The findings discussed here are preliminary and are based on the limited sample of the literature reviewed to date. Consequently, we are unable to report definitive results relating to every interaction effect within our moderators, at this stage. There is another methodological limitation caused by our limited sampling of the extant literature which is the range of confidence intervals in our results. We hope to alleviate this limitation, once we gather more effect sizes from our search.

4.2 Implications and Conclusion

Nonetheless, after initial findings, the AI technologies that exist in companies can be repurposed to put more emphasis on primary factors such as conversational, emotional, and explainable factors, which may increase individuals' acceptance and usage of these technologies. More importantly, prioritizing the perceptions and feeling relating to privacy protection while interacting with AI technologies is more crucial than other traditional antecedents of acceptance.

5 Appendix A

Category	Name	Definition	Coding	Variable Name
AI Characteristic	Used Acceptance Characteristic	Categorical variable representing which model does the characteristic used as a predictor variable belong to. (i.e. UTAUT, TAM, anthropomorphism, trust, risk, competency, satisfaction, enjoyment, external variables)	UTAUT, TAM, ANT, TRU, RISK, COMP, SAT, ENJ, EXT	CharType
Chatbot Relation	Relation Type	Categorical variable representing the intentional or actual relation to chatbot (e.g. usage intention, behavioral intention, attitude, satisfaction)	UsageInt, BehInt, Attitude, Satisfaction	RelType
Chatbot Interaction	Nature of Interaction	Categorical variable representing whether participant's interaction within the research is scenario-based or real. Dummy coded.	0 = Scenario-based 1 = Real interaction	ChatbotInterac
	Personal vs Professional Use	Categorical variable representing the personal or professional use of chatbot.	Pers Prof	PersvsProf
	Industry of Use	Categorical variable representing the industry that chatbot operating in (i.e. E-commerce, tourism, health, more than one industry)	E-commerce Tourism Health Complex	Industry
	Information vs Task	Categorical variable representing the nature of task chatbot carries out illustrated in the study (i.e. information giving, task completion)	Info Task	InfovsTask
	Conversation making	Categorical variable representing chatbot's conversational attributes in the study. Contrast coded.	-1 = No words 0 = No mention 1 = With words	Conv
Observable Chatbot Properties	Emotionality	Categorical variable representing chatbot's ability to carry out conversations that reflects emotionality via usage of words illustrated in the study for every participant. Contrast coded.	-1 = No emotional words used 0 = No mention 1 = Emotional words used	Emot
	Explainability	Categorical variable representing whether how the chatbot functions is explained to the participants. Dummy coded.	0 = Not explained 1 = Explained	Expl
Research Characteristics and Publication Bias	Year of publication	Continuous variable representing year that article is published		YearPub
	Sample size	Continuous variable representing the sample size used in the study of interest		n
	Number of measurement items	Continuous variable representing how many survey items was the outcome variable measured with.		ItemNo
	Sampling	Categorical variable representing whether the sample was a convenience sample (e.g. student). Dummy coded.	0 = Convenience sample 1 = Random sample	SampleType
	Manipulation	Continuous variable representing the number of variables manipulated within the study of interest.		VarManip
	Data collection location	Categorical variable representing the country of data collection in the study		DataLoc
	Mean age	Continuous variable representing the mean age of participants		Mean_age
	Gender (Female)	Continuous variable representing the percentage of females to the whole sample in the study		Fem_perc
	Study Design	Categorical variable representing whether the study design was within- or between-subject, or cross-sectional		Design
	Empirical Observation	Categorical variable representing the empirical nature of the design (i.e. experiment, correlational, N/A)		ExpDesign
	Publication Status	Categorical variable representing whether the article was published with peer-review. Dummy coded.	0 = Not published 1 = Published	PubStatus
	Effect size Precision	Continuous variable indicating the precision of the effect size, calculated by the inverse of variability of *z* scores		Precision
	Zero-order vs Partial-Correlation	Categorical variable representing whether the effect size is extracted from a correlation matrix or a statistical test	0 = Zero-order 1 = Partial correlation	ZeroParCorr

Fig. 2. Coding scheme followed for 18 articles in the sample.

6 Appendix B

For the next stage of our scoping, we have created an inclusion criteria list following these terms:

1. Article of interest must be published or printed within a decade, between 2012–2021.
2. These articles should be either from a peer-reviewed journal or from gray literature.

3. The article should address the research questions we have indicated for this meta-analysis:
   (a) The article must investigate technology acceptance models or other characteristics as an independent or predictor variable.
   (b) The article must investigate the level of acceptance, which may be conceptualized as usage or adoption intentions, attitudes, or satisfaction as a dependent or observed variable.
4. In each article of interest, results of statistical analysis models should indicate sufficient statistics (t- or p-value, sample size, etc.) to extract an effect size in the form of Cohen's d or r.
5. The sample cannot include any research employed to vulnerable groups (i.e. populations who cannot consent to participate or individuals under 18 years of age).
6. Article must be written in English.
7. Article must have full-text availability.
8. If the data used in an article was also used in another publication, only one article must be selected.

References

1. Croes, E.A.J., Antheunis, M.L.: Can we be friends with Mitsuku? a longitudinal study on the process of relationship formation between humans and a social chatbot. J. Soc. Pers. Relat. **38**(1), 279–300 (2020). https://doi.org/10.1177/0265407520959463
2. Davis, F.: A technology acceptance model for empirically testing new end-user information systems: theory and results. Massachusetts Institute of Technology (1986). http://hdl.handle.net/1721.1/15192
3. Egger, M., Smith, G.D., Schneider, M., Minder, C.: Bias in meta-analysis detected by a simple, graphical test. BMJ **315**(7109), 629–634 (1997). https://doi.org/10.1136/BMJ.315.7109.629
4. Følstad, A., Nordheim, C.B., Bjørkli, C.A.: What makes users trust a chatbot for customer service? an exploratory interview study. In: Bodrunova, S.S. (ed.) INSCI 2018. LNCS, vol. 11193, pp. 194–208. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01437-7_16
5. Huang, D.H., Chueh, H.E.: Chatbot usage intention analysis: veterinary consultation". J. Innov. Knowl. **6**(3), 135–144 (2021). https://doi.org/10.1016/J.JIK.2020.09.002
6. Hudon, A., Demazure, T., Karran, A., Léger, P.-M., Sénécal, S.: Explainable artificial intelligence (XAI): how the visualization of ai predictions affects user cognitive load and confidence. In: Davis, F.D., Riedl, R., vom Brocke, J., Léger, P.-M., Randolph, A.B., Müller-Putz, G. (eds.) NeuroIS 2021. LNISO, vol. 52, pp. 237–246. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-88900-5_27
7. Ischen, C., Araujo, T., Voorveld, H., van Noort, G., Smit, E.: Privacy concerns in chatbot interactions. In: Følstad, A., Araujo, T., Papadopoulos, S., Law, E.L.-C., Granmo, O.-C., Luger, E., Brandtzaeg, P.B. (eds.) CONVERSATIONS 2019. LNCS, vol. 11970, pp. 34–48. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-39540-7_3

8. Kaplan, A.D., Kessler, T.T., Brill J.C., Hancock, P.A.: Trust in artificial intelligence: meta-analytic findings. Human Factors (2021). https://doi.org/10.1177/00187208211013988
9. Qiu, L., Benbasat, I.: Evaluating anthropomorphic product recommendation agents: a social relationship perspective to designing information systems, vol. 25, no. 4, pp. 145–182, April 2014, https://doi.org/10.2753/MIS0742-1222250405
10. Rosenthal, R.: The file drawer problem and tolerance for null results. Psychol. Bull. **86**(3), 638–641 (1979). https://doi.org/10.1037/0033-2909.86.3.638
11. Sheehan, B., JinH. S., Gottlieb, U.: Customer service chatbots: anthropomorphism and adoption. J. Bus. Res. **115**, 14–24 (2020). https://doi.org/10.1016/J.JBUSRES.2020.04.030
12. Duval, S., Tweedie, R.: Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics **56**(2), 455–463 (2000). https://doi.org/10.1111/J.0006-341X.2000.00455.X
13. Venkatesh, V., Thong, J.Y.L., Xu, X.: Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. MIS Quarterly: Manage. Inf. Syst. **36**(1), 157–178 (2012). https://doi.org/10.2307/41410412
14. Viechtbauer, W.: Conducting meta-analyses in R with the metafor package. J. Statist. Softw. **36**(3), 1–48, (2010). 10.18637/jss.v036.i03
15. Xu, J., Benbasat, I., Centefelli, R.T.: The nature and consequences of trade-off transparency in the context of recommendation agents. MIS Q. **38**(2), 379–406 (2014)