Aug 10th, 12:00 AM

Adoption of Smart Home Speakers: Study of User Acceptance Factors

Evgeniia Filippova-Karlusch
MCI Management Center Innsbruck, evgeniia.filippova@mci.edu

Jakob Vincent Summer
Capgemini Invent, jakob.summer@freenet.de

Adoption of Smart Home Speakers: Study of User Acceptance Factors

Completed Research

Evgeniia Filippova-Karlusch
MCI Management Center Innsbruck
evgeniia.filippova@mci.edu

Jakob Vincent Summer
Capgemini Invent
jakob.summer@capgemini.com

Abstract

Smart speakers have entered the market in recent years and have already been adopted by certain parts of the population. Although several studies have previously tried to explain some aspects of smart speaker adoption, there is still no common understanding in the literature about the crucial factors affecting their acceptance. Building on an extensive analysis of innovation adoption theories with a special focus on studies of voice-controlled technologies, the authors cover this research gap by constructing a research model for user acceptance factors of smart speakers and testing it empirically within a sample of current non-adopters in Germany. The results show that enjoyment, privacy concerns, usefulness and social influence may significantly predict the intention to use smart speakers. Moreover, system quality and complementarity of smart speakers with various apps and services are predictors of perceived usefulness.

Keywords

Smart speaker, voice assistance, adoption, user acceptance.

Introduction

Smart speakers are on the rise. Since the release of one of the first voice-empowered home speakers, the Amazon Echo in 2014 sales have increased rapidly. At the beginning of 2020, reportedly one-third of households in the United States owned at least one smart speaker. However, despite its significant initial growth, the penetration rate of smart speakers, compared to other household IT devices such as mobile phones, notebooks, and connected televisions, is still relatively low, and saturation is far from in sight (Fang and Fu 2020).

As companies could benefit significantly from the smart speaker market, not only and not so much through high profits from the sales of the devices themselves but especially from gaining valuable customer data to improve speech recognition algorithms and other offerings (Cheng et al. 2019), it is of crucial importance for them to ensure a broad adoption of smart speakers. Given the great potential of the smart speaker market and its still relatively low adoption rate (Fang and Fu 2020), the following question arises*: what user acceptance factors, and to which extent, affect the intention to use smart speakers among the current non-adopters?*

Although there have been several authors who investigated the adoption of smart speakers and their voice empowered software agents over the past years (Ling et al. 2021), there is still a lack of common understanding of the factors affecting the intention to use these devices at their current technology level. To cover this research gap, the authors first develop a research model based on the extensive meta-analysis of 151 IT innovation adoption papers and a review of 22 smart speaker adoption studies conducted across different countries. Subsequently, the research model is empirically tested within the sample of adults living in Germany who are active internet users but do not yet own or regularly use a smart speaker.

The paper is structured as follows. Section 2 sheds light on the current literature on voice-enabled technologies and provides an overview of the innovation adoption theories. The empirical findings of previous IT and smart speaker adoption studies are also summarized in the second section. Based on the

extensive literature analysis, the research model and corresponding hypotheses are constructed in Section 3 of the present paper. Employed methods for the empirical investigation are shortly described in Section 4, whereas the empirical findings are the subject of Section 5. Section 6 concludes and suggests possible directions for further research.

Theoretical background

Voice operated technology

An Intelligent Personal Assistant (IPA) also referred to as a "digital personal assistant" (Lopatovska 2019) or a "smart personal assistant" is an application that evaluates user inputs such as voice, images, text, or contextual information and provides an appropriate response, for example in the form of an answer or the execution of a task (Lopatovska 2019). While some authors argue that all IPAs can interpret spoken language, others consider this ability as a common, however, not necessary feature of an IPA. For the purposes of the present research, the authors refer to a **voice assistant** as an intelligent personal assistant that is capable of (i) interpreting natural spoken language and possible also other user input, and (ii) responding via artificial voice and other means (Hoy 2018). Other terms that are found in the current literature to be interchangeably used for the voice assistants include "voice-activated personal assistants", "voice-activated digital assistants", "voice-controlled agents", "conversational agents", "virtual voice assistants", and others (Lopatovska 2019; Koon et al. 2020).

Cloud-based voice assistants like Siri from Apple or Google Assistant from Google are widely spread on mobile devices and are powered by a combination of several digital technologies, such as the Internet of Things, cloud computing, and natural language processing – a branch of AI that enables the software agent to automatically process natural spoken language, encode it into structured information and respond to it. The functionalities of voice assistants increased rapidly over the past years and can be summarized as dialogue, web search and chat functions (Jiang et al. 2015). The dialogue function allows users to interact with the device via voice to carry out specific tasks or access certain features. The web search function enables users to search for information online via voice. The chat function enables users to interact with the voice assistant and receive predefined responses. Among others, voice assistants can send text messages, make phone calls, set reminders, tell jokes, play music, or control other smart home devices such as lighting systems or thermostats (Hoy 2018). Some voice assistants can be individually updated with additional functions, so-called voice apps. Voice apps for Amazon's Alexa are called skills. Back in 2019 there were already over 100,000 skills available for Alexa-empowered smart speakers: examples include 'Domino's' to order pizza, 'Spotify' to access Spotify accounts and play music, and 'PayPal' to make financial transactions via voice.

Lacking a unified definition, a **smart speaker** is commonly referred to as a stand-alone device with integrated microphones and speakers as main hardware components that can be connected via interfaces to a voice assistant (Kowalczuk 2018). Smart speakers might vary in size, design, sound quality, supported voice assistants, and provision of visual feedback in the form of lights. At the end of 2020, the most sold smart speakers globally were those from Amazon, Google, Baidu, Alibaba, Apple, and Xiaomi. In Germany the first smart speaker – Amazon Echo – was launched in 2016, and four years later reportedly around 24% of the adult population owned at least one smart speaker. These findings allow to categorise with caution the current users of the smart speakers in Germany as the early majority on the innovation adoption curve of Rogers.

IT innovation adoption

There is no unified theory of innovation adoption, nor it is likely to be developed. Instead, multiple theories and theoretical frameworks describe factors that have an impact on the adoption of IT innovations (Hameed et al. 2012). The well-established theories in the field, such as Diffusion of Innovation (DOI), Technology Acceptance Model (TAM), and the Unified Theory of Acceptance and Use of Technology (UTAUT) among others, differ in the influencing factors considered and have their specific advantages and limitations. Since the mentioned theories – together with additional frameworks from the IT innovation adoption studies – lay a theoretical foundation for the subsequent empirical investigation within the present study, their main points are shortly outlined in the next paragraphs.

The Diffusion of Innovation Theory developed by E. M. Rogers in 1962 represents the oldest theory in the field. Within this theory, there have been identified five perceived characteristics of an innovation that affect the attitude towards it: (i) *relative advantage* of the innovation in focus; (ii) *compatibility* with existing values, earlier experiences and needs of potential adopters; (iii) *complexity* in terms of understanding and usage of innovation; (iv) *trialability*; and (v) *observability* addressing the visibility of the results of innovation to others (Rogers 2003). Representing undoubtedly one of the foundational frameworks in the innovation adoption literature, the perceived characteristics are often criticized because of generalization and, therefore, limited explanatory power for the adoption of specific types of innovations, such as IT innovations (Momani et al. 2017). Subsequently, there have been numerous attempts to extend the five perceived characteristics of innovation, adding communicability, social approval, costs, voluntariness, image, and others (Tornatzky and Klein 1982; Moore and Benbasat 1991).

Technology acceptance models in general – Theory of Reasoned Action (TRA), Theory of Planned Behavior (TPB), and Technology Acceptance Model (TAM) – aim to explain human acceptance factors that affect the intention towards using a certain technology. While the TRA and the TPB rather explain factors influencing human behavior in general, the TAM has been originally developed to predict the acceptance and usage of IT by individuals at their workplace (Davis 1989). According to the TAM, two constructs – *perceived usefulness* and *perceived ease of use* – have a primary effect on users' behavioral intention to use IT innovations and their subsequent usage. Both factors are also affected by external variables. Additionally, it was found that an IT system is perceived as more useful if it is easy to use (Davis 1989). Being one of the most used frameworks in the empirical IT innovation adoption studies so far, TAM is being criticized for oversimplicity due to a small number of variables considered, and a rather poor explanatory power in people's self-reported system use (Legris et al. 2003). To address the weak points of the model, it was extended to TAM2 (Venkatesh and Davis 2000) and TAM3, as well as to UTAUT (Venkatesh et al. 2003) and UTAUT2 to include additional external variables that affect the behavioral intention to use IT in different settings. While TAM 1-3 and UTAUT were elaborated to determine factors that affect the acceptance of IT by individuals within an organization, the subsequent version of UTAUT2 was developed to study the acceptance of IT in a consumer context.

The Unified Theory of Acceptance and Use of Technology 2 (UTAUT2) is based on the former technology acceptance models and includes seven factors that affect the behavioural intention and subsequent usage of IT by individuals: (i) *performance expectancy* addressing the benefits to consumers; (ii) *effort expectancy* which is similar to the previous ease of use; (iii) *social influence*; (iv) *facilitating conditions* reflecting the resources available to use the technology; (v) *hedonic motivation* which represents the fun and pleasure from technology usage; (vi) *price value*; and (vii) *habit*. Furthermore, UTAUT2 considers three moderating variables – age, gender, and experience. Being a widely used model to analyse the acceptance and usage of IT within a consumer context, the UTAUT2 model is commonly complemented with additional factors.

In addition to the well-established innovation adoption theories previously described, there are numerous other frameworks in place explaining relevant adoption factors. For example, the study of Hameed et al. (2012) provides an overview of user acceptance factors considered in previous IT adoption literature and their corresponding significance for acceptance and actual usage of IT innovations at the individual level. The significant factors, among others, include perceived voluntariness, anxiety, compatibility, perceived enjoyment, user training, facilitating conditions, financial incentives, technical assistance, subjective norms, self-efficacy, and others (Hameed et al. 2012). Based on the extensive literature review, the authors identified a range of user acceptance factors for smart speaker adoption which are shortly summarized below.

User acceptance factors for adoption of smart speakers

To identify the user acceptance factors for the adoption of smart speakers in the existing literature, the following study selection criteria have been applied: (i) the study investigates factors influencing the adoption, acceptance or (continuous) usage of voice assistants (accessible via smart speaker) or smart speakers as such; and (ii) the study analyzed the acceptance or use of smart speakers by individuals and not organizations. Peer-reviewed articles were preferred, but other publications have been considered as well. In total, twenty-two studies of quantitative and qualitative nature have been identified and thoroughly analyzed. As a result, the factors identified were clustered into 19 categories: usefulness, usability,

enjoyment, privacy/security concerns and trust, social influence, self-representation, personal relationship with device, facilitating conditions, system/platform related variables, visual attraction, brand, monetary aspects, value, attitude, compatibility, portability, habit, job relevance, and relative advantage. The factors that proved to have a significant effect on the intention of the usage of smart speakers are shortly explained in the paragraphs below.

Usefulness relates to the perceived utility of smart speakers, mainly referring to the consumer perspective of 'perceived usefulness' construct of the TAM. Several authors relate to usefulness as performance expectancy, convenience, task attraction or utilitarian benefits (Han and Yang 2018; Koon et al. 2020). *Usability* refers to all constructs related to the perceived ease of using smart speakers and is reported to have a positive effect on the intention to use them. *Enjoyment* includes constructs related to the perceived emotional value, hedonic motivation or fun expected from using smart speakers (Ling et al. 2021). As privacy considerations might retain individuals from using smart speaker devices (Han and Yang 2018), *privacy/security concerns* have been analyzed by several authors to have a negative effect on the acceptance of smart speakers. *Social influence* includes factors concerning the impact that social surrounding has on an individual's intention towards using smart speakers. Some researchers identified a positive effect of social influence on the adoption intention, while others did not (Coskun-Setirek and Mardikyan 2017). For certain smart speakers, such as Alexa, third-party providers can create voice apps or complementary products that can be controlled via the voice assistant. Therefore, different constructs related to the corresponding ecosystem of voice assistants' platforms (*system and platform related factors*) have been considered as having an effect on the acceptance of smart speakers in previous studies (Coskun-Setirek and Mardikyan 2017). *Personal relationship with the device* refers to the social bond people feel toward smart speakers and includes constructs like anthropomorphism, social presence, and para-social relationship (Han and Yang 2018; Moussawi et al. 2020). Previous authors found a positive effect of personal relationship with the device on the adoption intention. The *Self-representation* cluster refers to image considerations that might positively affect the usage intention of smart speakers (McLean and Osei-Frimpong 2019). *Monetary aspects*, such as price or perceived fee, have also been considered relevant factors in the previous studies with controversial empirical evidence (Park et al. 2018; Ling et al. 2021). Lastly, a *visual attraction* or a *perceived design attractiveness* of smart speakers is said to have an indirect positive effect on an individual's intention to use the device, moreover, because smart speakers are often placed in rooms where people tend to spend more time, such as bedrooms or living rooms.

Research model and hypotheses development

Based on the analysis of user acceptance factors of IT innovation adoption in general and smart speakers in particular, the following research model for subsequent empirical investigation has been developed.

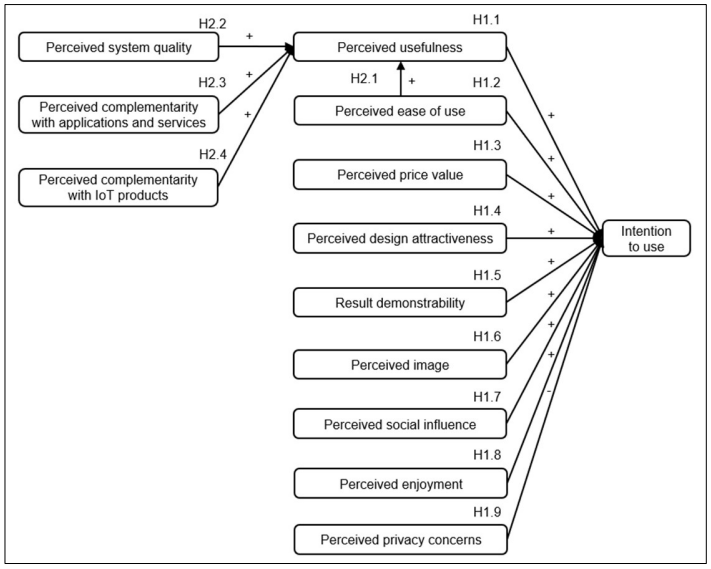

Figure 1. Research Model

Hypotheses on factors affecting the usage intention

In line with the Technology Acceptance Model (TAM), perceived usefulness and perceived ease of use have a positive effect on the intention to use IT innovations (Davis 1989). Taking in consideration the findings of the previous studies on smart speakers' adoption (Han and Yang 2018; Koon et al. 2020) and other IT innovation in general, it is assumed that both the perceived usefulness and perceived ease of use positively affect the intention to use smart speakers (*H1.1* and *H1.2* correspondingly). The perceived price value refers to the tradeoff between the cost of the device and the benefits of using it. In case of a positive perceived price value, the benefits are expected to outreach the monetary costs, therefore it is assumed that perceived price value will have a positive effect on the intention to use smart speakers (*H1.3*). According to the previous studies, perceived design attractiveness plays an important role in the adoption decision of smartphones, smart watches, and other similar devices (Hsiao 2017). Considering the findings of the previous studies combined with the fact that smart speakers are often placed in the rooms where people tend to spend more time, such as bedrooms or living rooms, it is assumed that perceived design attractiveness positively affects the intention to use smart speakers (*H1.4*).

Based on the results of the existing IT adoption studies, it is also assumed that result demonstrability – like the observability construct of Rogers (Moore and Benbasat 1991) – has a positive effect on smart speakers' usage intention (*H1.5*). Perceived image – a construct mentioned in the TAM3 and other studies investigating individual adoption of innovative digital devices, such as smart glasses (Rauschnabel et al. 2018) – is assumed to have a positive effect on the intention to use smart speakers as well (*H1.6*). Perceived social influence – "the degree to which an individual perceives that important others believe he or she should use [a certain technology]" (Venkatesh et al. 2003) – has been commonly identified as a factor that positively affects the intention to use technology and IT innovations. This leads to the next hypothesis, that perceived social influence positively affects the intention to use smart speakers (*H1.7*). Perceived enjoyment, being a commonly addressed factor that might have a positive effect on the usage of IT devices with enjoyment components, is assumed to positively affect the intention to use smart speakers as well (*H1.8*). Previous studies identified that perceived privacy concerns toward the use of certain IT services and IoT devices might negatively affect the adoption intention of the users (Hsu and Lin 2018). The privacy considerations apply to smart speakers as well. As voice assistants are activated via a keyword, they are in constant listening mode. Once activated they gather the users' data, which are processed on the service provider's cloud to produce appropriate responses. Conversational recordings are also stored and used by the service providers to enhance their services and offerings. This leads to a potential risk of data misuse; therefore, perceived privacy concerns are assumed to negatively affect the intention to use smart speakers (*H1.9*).

Hypotheses on factors affecting perceived usefulness

The research model of the present study takes into consideration additional factors that affect the perceived usefulness of smart speakers. The original TAM model of Davis, as well as its subsequent extensions, proved the positive effect of the perceived ease of use on the perceived usefulness of innovations, which is assumed to hold true also in the context of smart speakers (*H2.1*). Perceived system quality relates to the degree to which IT systems performs their functions and can be determined through reliability, response time, or the capability to combine information from multiple sources. With consideration of the findings of previous studies on smart speakers, it can be assumed that perceived system quality would positively affect the perceived usefulness of smart speakers (*H2.2*). Another important factor for IT innovations is perceived complementarity (Hsu and Lin 2018). The research model within the present study differentiates between perceived complementarity of smart speakers with applications and services, and with hardware IoT components. Given the vast availability of voice apps for smart speakers, as well as increasing convergence between these devices and other smart home systems, for example lightning, locks or radiators, both of these complementarities are assumed to have a positive effect on the perceived usefulness and represent the *H2.3* and *H2.4* correspondingly.

Method

In line with most previous studies on the subject, the authors used a quantitative approach within this paper as well. An online survey was conducted with a sample (n=233) of adults living in Germany who are active internet users and do not own or regularly use smart speakers yet. Non-internet users were excluded from the target population, as an internet connection is a prerequisite for the usage of a smart speaker. The focus on the people who do not own or regularly use smart speakers can be explained by several reasons. First, predictors for the usage intention should be measured prior to or with the decision to (initially) use an innovation (Tornatzky and Klein 1982). Retrospective user acceptance studies might be biased, as people who have already adopted an innovation tend to rationalize their decision by rating the characteristics of this innovation more favorably than initially perceived. Second, as smart speakers are constantly improving, current adopters might only be able to express their perception of older and potentially outdated smart speaker models.

To recruit participants for the survey, a non-probability sampling technique, which implies certain limitations to the generalizability of the study mentioned in the last section of the paper, was applied. In total, 327 respondents participated in the web survey that was distributed via multiple channels between May and June 2021. Incomplete questionnaires or questionnaires filled in by respondents not matching the target population were sorted out prior to the analysis. The final sample consisted of 233 participants. The survey consisted of in total 36 items to measure the 13 constructs of the model. Additionally, questions on the socio-economic characteristics of respondents, such as age, gender, country of residence, average annual household net income and highest completed level of education were designed and placed at the end of survey.

To guarantee a certain degree of cross-study compatibility, the measurement items for the thirteen constructs of the research model were all adapted from previous adoption studies. The original wording was slightly adapted to smart speakers and their specific characteristics. The items to measure the perceived usefulness and perceived ease of use were adapted from Davis (1989). Perceived image and result demonstrability were measured using the items developed by Moore and Benbasat (1991). The slightly adapted items of Rauschnabel et al. (2018) to measure perceived privacy concerns. The items to measure perceived complementarity with applications and services, as well as perceived complementarity with IoT products were adapted from Hsu and Lin (2018). To measure the construct perceived system quality, items of Barki and Hartwick (2001) were used. To ensure consistency, all questions were measured on an ordinal seven-point Likert scale with the terms ‘strongly disagree’ and ‘strongly agree’ as the minimum and maximum values. Due to the limited space a detailed table with the items to measure each construct and the wording of the scale is not provided.

A multiple linear regression analysis (MLRA) was applied to analyze the effects of the independent variables on the dependent variables. MLRAs are designed to analyze and predict the effect (strength and direction) between a dependent metrically scaled variable and several categorial or metrically scaled independent variables that are hypothesized to affect it. MLRA can also assess the unique effect of each single predictor within the set on the independent variables. The findings of the MLRA do not allow causal interferences but do enable to determine how a set of independent variables is associated with the dependent variable of interest (Aiken et al. 2013).

After the filtering process and the import of the data into SPSS, the data analysis was conducted. First, descriptive statistics of the survey results were created. Next, the internal reliability of the scales was tested using Cronbach’s alpha (α). After ensuring that the items used to measure each construct correlated sufficiently strongly among each other to be combined into one scale, descriptive information about the constructs was summarized. Then, the two separate MLRAs to analyze the hypothesized effects of the user acceptance factors within the research model were conducted. The required assumptions for a MLRA – no multicollinearity, no autocorrelation, normal distribution of residuals, and homoscedasticity – have been statistically tested. No-multicollinearity was tested by calculating and analyzing the variance inflation factor (VIF) – since all VIF values were below ten (ranging from 1.10 to 1.40), the assumption was satisfied (Aiken et al. 2013).

Findings

Descriptive statistics of the survey results

The analysis of the socio-economic characteristics of the sample revealed that the respondents were predominantly under 30 years old and most of them had university degrees. Regarding the interest in various smart speaker functions, above half of the respondents were interested in streaming music, setting timers and alarms, and controlling smart home devices with the use of voice. About 27% of the respondents were interested in receiving calendar updates, conducting web searches, and sending text messages. Only a small percentage of the respondents were interested in the voice shopping function and the possibility to play games via smart speakers. Eighteen percent stated that they were not interested in any of the mentioned smart speaker functions. In such a way, the respondents showed especially high interest in the basic functions of smart speaker devices (streaming musing or setting timers), while the interest in more advanced features was substantially lower.

The awareness of the respondents of the smart speaker brands revealed the following insights: the smart speakers of Amazon were the most well-known among the participants, followed by Apple, Google and Sonos. Four percent of the sample were not aware of any of the mentioned smart speaker brands. Furthermore, the selection option 'others' was placed as an open question and the following two smart speaker manufacturers were stated once: Lenovo and Ultimate Ears. These findings are in line with a previous study conducted by Statista back in 2019, which also identified that Germans were most interested in smart speakers from Amazon, Apple, and Google. Furthermore, the respondents stated on average relatively high privacy concerns towards using smart speakers. One possible reason for this might be the frequent news reports in Germany about a lack of transparency and data protection from smart speaker providers, which may cause current non-adopters – who in general are more risk averse – to be concerned about their data privacy. The low mean of perceived image and perceived social influence might indicate that the respondents, on average, did not perceive smart speakers as a status symbol and were generally not encouraged by their social environment to use them.

Effects on the perceived usefulness of smart speakers

A MLRA1 was calculated to test if perceived usefulness of smart speakers can be predicted by the four hypothesized predictors – perceived ease of use (PEOU), perceived system quality (PSQ), perceived complementarity with applications and services (PCAS), and perceived complementarity with IoT products (PCIP). The adjusted R^2 value was 0.176, therefore the perceived usefulness could be explained by 17.6% from its hypothesized predictors. The F-test showed a score of $F(4,228) = 13.38$, $p<0.001$, therefore the regression model as a whole has significant explanatory power for predicting perceived usefulness. The perceived system quality has a significant positive effect on perceived usefulness. Also, a significant and the largest positive effect of perceived complementarity with applications and services on perceived usefulness of smart speakers could be identified. For the other independent variables – perceived ease of use and perceived complementarity with IoT products – no significant effect on perceived usefulness could be identified. The regression analysis summary is provided in the Table 1.

Variable	*B*	*SE*	*Beta*	*t*	*p*
Constant	-0.20	0.61		-0.32	0.751
PSQ	0.29	0.10	0.18	2.80	0.006
PEOU	0.08	0.07	0.06	1.02	0.309
PCAS	0.40	0.11	0.25	3.56	<0.001
PCIP	0.14	0.10	0.09	1.40	0.164

Table 1. Regression Analysis Summary for MLRA1

Effects on the usage intention of smart speakers

A second MLRA was calculated to test, whether behavioral intention to use smart speakers could be predicted from the nine independent variables: perceived usefulness (PU), perceived ease of use (PEOU), perceived price value (PPV), perceived design attractiveness (PDA), result demonstrability (RD), perceived image (PI), perceived social influence (PSI), perceived enjoyment (PE), and perceived privacy concerns (PPC). The adjusted R^2 value indicates that behavioral intention could be explained by 55.4% from its

hypothesized predictors. The F-test showed a score of F(9,223) = 33.00, p<0.001, therefore the regression model as a whole has significant explanatory power for predicting behavioural intention.

Perceived usefulness was found to have a significant positive effect on the intention to use smart speakers. Also, a significant positive effect of perceived enjoyment and perceived social influence on the usage intention could be identified. A significant but negative effect of the privacy concerns could be detected. For the other independent variables, no significant effect could be identified. The summary of the regression results is provided in the Table 2 below.

Variable	*B*	*SE*	*Beta*	*t*	*p*
Constant	1.84	0.69		2.65	0.009
PU	0.26	0.06	0.22	4.08	<0.001
PE	0.43	0.06	0.37	6.96	<0.001
PDA	0.03	0.07	0.02	0.37	0.712
PPC	-0.36	0.07	-0.27	-5.51	<0.001
PI	0.01	0.07	0.10	0.18	0.859
PSI	0.20	0.07	0.16	2.88	0.004
PPV	0.07	0.08	0.04	0.83	0.408
PEOU	0.03	0.07	0.03	0.48	0.634
RD	-0.04	0.08	-0.03	-0.49	0.623

Table 2. Regression Analysis Summary for MLRA2

Hypotheses validation and revised research model

To sum up, six out of the thirteen developed hypotheses were supported through the empirical study. The results are visualized in the Figure 2 below, where the rejected hypotheses are shown with dashed lines.

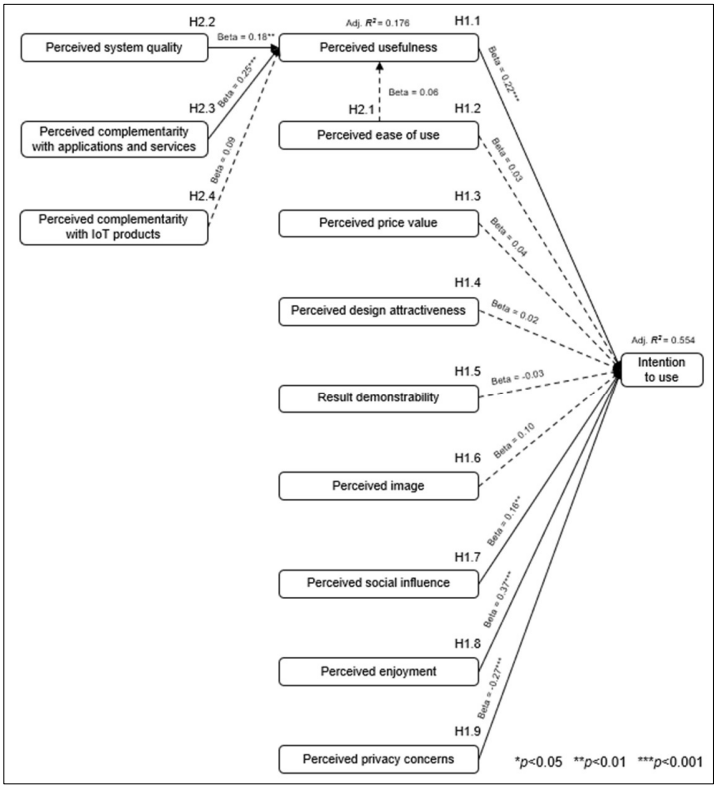

Figure 2. Revised Research Model

As visible in the Figure 2, the hypotheses on the positive effects of perceived system quality and perceived complementarity with applications and services (*H2.2* and *H2.3* accordingly) were confirmed. This result confirms the findings of prior researchers who identified similar effects. Moreover, the significant positive effect of perceived complementarity with applications and services on perceived usefulness supports the theory of (indirect) network externalities. The positive effect of perceived ease of use on perceived

usefulness of smart speakers in line with Davis (1989) could not be identified within the present study. A possible explanation could be an already mature usability of smart speakers, which is expected from all the devices and does not necessarily lead to a higher perceived usefulness.

The study revealed positive effects of perceived usefulness, perceived enjoyment, and perceived social influence on the behavioral intention to use smart speakers (*H1.1*, *H1.8*, and *H1.7* accordingly). The results on perceived enjoyment and perceived usefulness supports the findings of previous studies in the domain (Kowalczuk 2018). Furthermore, the positive effect of the perceived usefulness is also in line with the TAM3 and UTAUT2 models. The study at hand showed that perceived privacy concerns might have a significant negative effect on the individuals' intention to use smart speakers (*H1.9*). This goes in line with the other technology acceptance studies in general. Unlike some other studies in the field of smart speaker acceptance, no other positive effects of the rest of the identified factors could be detected. For example, no positive effect of perceived design attractiveness on the intention to use smart speakers could be detected. A possible reason for this might be that this variable rather affects the perceived value of the device as suggested by Rindova and Petkova (2007) for the novel product in general, therefore, having only an indirect effect on behavioral intention. Another possible explanation might be that the design of IT products in general has already become so mature and dominant that an appealing design itself might no longer be regarded as a factor that significantly positively affects intention to use.

Conclusion and outlook

The study at hand identified user acceptance factors that affect the intention to use smart speakers among current non-adopters in Germany. The authors made several contributions to the existing literature in the domain. First, the analyzed user acceptance factors were derived based on an extensive literature review of previous smart speaker adoption studies and commonly applied theoretical framework. Second, the study – to the best of the authors' knowledge – is the first one to study the acceptance of smart speakers at the current technology state of the devices and on the current state of the adoption curve in Germany. Different to other papers in the field, the study at hand analyzed what affects the intention to use smart speakers among the later adopter categories. Third, this study considered a relatively comprehensive number of user acceptance factors in its underlying research model and additionally investigated the factors affecting not only the intention to use, but also the perceived usefulness of the devices. The results of the study might be of interest to practitioners as well, since they provide valuable indicators on which factors smart speaker manufacturers might focus to further promote the acceptance and consequently the actual usage of these devices in Germany. In fact, producers might focus particularly on improving and marketing the enjoyment, usefulness, data security, and social aspects of smart speakers.

The study comes with several limitations. Since a non-probability sampling technique has been applied, and the sample size is relatively small, it is difficult to generalize results for the whole target population. The findings of the factors affecting the acceptance of smart speakers are only valid for the current technology stage of smart speakers. A similar study might be necessary as smart speaker adoption progresses to later stages on the adoption curve. Moreover, the fact that the survey was conducted among German respondents poses a limitation as well. It is likely that conducting a similar analysis in other countries would reveal different results based on, for example, cultural differences – this represents a promising avenue for the further research. It would be especially interesting to compare the impact of perceived privacy concerns of countries with different cultural backgrounds. The subsequent studies in the field should also explicitly consider the increasingly important ethical considerations of smart speaker usage, such as data ownership issues of vast amount of user voice data, data security – which is an increasing concern of cloud-based applications in general, and possibly cyber bullying threats.

REFERENCES

Aiken, L. S., West, S. G., Pitts, S. C., Baraldi, A. N., and Wurpts, I. C. 2013. "Multiple Linear Regression," In *Handbook of psychology*, pp. 511-542.

Barki, H., and Hartwick, J. 2001. "Interpersonal Conflict and Its Management in Information System Development," *MIS Quarterly* (25:2), pp. 195-228.

Cheng, P., Bagci, I. E., Yan, J., and Roedig, U. 2019. "Smart Speaker Privacy Control - Acoustic Tagging for Personal Voice Assistants," *2019 IEEE Security and Privacy Workshops (SPW)*, IEEE, pp. 144-149.

Coskun-Setirek, A., and Mardikyan, S. 2017. "Understanding the Adoption of Voice Activated Personal Assistants", *International Journal of E-Services and Mobile Applications* (9:3), pp. 1-21.

Davis, F. D. 1989. "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS quarterly* (13:3), pp. 319-340.

Fang, T., and Fu, X. 2020. „Development Status and Marketing Strategy of Smart Speakers," In *International Conference on Applied Human Factors and Ergonomics*, Springer, Cham, pp. 553-562.

Hameed, M. A., Counsell, S., and Swift, S. 2012. "A Conceptual Model for the Process of IT Innovation Adoption in Organizations," *Journal of Engineering and Technology Management* (29:3), pp. 358-390.

Han, S., and Yang, H. 2018. "Understanding Adoption of Intelligent Personal Assistants: A Parasocial Relationship Perspective," *Industrial Management & Data Systems* (118:3), pp. 618-636.

Hoy, M. B. 2018. "Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants," *Medical reference services quarterly* (37:1), pp. 81-88.

Hsiao, K.-L. 2017. "What Drives Smartwatch Adoption Intention? Comparing Apple and non-Apple Watches," *Library Hi Tech* (35:1), pp. 186-206.

Hsu, C.-L., and Lin, J. C.-C. 2016. "An Empirical Examination of Consumer Adoption of Internet of Things Services: Network Externalities and Concern for Information Privacy Perspectives", *Computers in Human Behavior* (62), pp. 516–527.

Jiang, J., Hassan Awadallah, A., Jones, R., Ozertem, U., Zitouni, I., Gurunath Kulkarni, R., and Khan, O. Z. 2015. "Automatic Online Evaluation of Intelligent Assistants", In *Proceedings of the 24th International Conference on World Wide Web*, pp. 506-516.

Koon, L. M., McGlynn, S. A., Blocker, K. A., and Rogers, W. A. 2020. "Perceptions of Digital Assistants from Early Adopters Aged 55+," *Ergonomics in design* (28:1), pp. 16-23.

Kowalczuk, P. 2018. "Consumer Acceptance of Smart Speakers: A Mixed Methods Approach", *Journal of Research in Interactive Marketing* (12:4), pp. 418-431.

Legris, P., Ingham, J., and Collerette, P. 2003. "Why Do People Use Information Technology? A Critical Review of the Technology Acceptance Model," *Information & Management* (40:3), pp. 191–204.

Ling, H. C., Chen, H. R., Ho, K. K., and Hsiao, K. L. 2021. "Exploring the Factors Affecting Customers' Intention to Purchase a Smart Speaker," *Journal of Retailing and Consumer Services* (59), 102331.

Lopatovska, I. 2019. "Overview of the Intelligent Personal Assistants", *Ukrainian Journal on Library and Information Science*, (3:72), pp. 72-79.

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.

Momani, A. M., Jamous, M. M., and Hilles, S. M. 2017. "Technology Acceptance Theories: Review and Classification", *International Journal of Cyber Behavior, Psychology and Learning* (7:2), pp. 1-14.

Moore, G. C., and Benbasat, I. 1991. "Development of an Instrument to Measure the Perceptions of Adopting an Information Technology Innovation," *Information systems research* (2:3), pp. 192-222.

Moussawi, S., Koufaris, M., and Benbunan-Fich, R. 2020. "How Perceptions of Intelligence and Anthropomorphism Affect Adoption of Personal Intelligent Agents", *Electronic Markets* , pp. 343-364.

Park, K., Kwak, C., Lee, J., and Ahn, J. H. 2018. "The Effect of Platform Characteristics on the Adoption of Smart Speakers: Empirical evidence in South Korea", *Telematics and Informatics* (35:8), pp. 2118-2132.

Rauschnabel, P. A., He, J., and Ro, Y. K. 2018. "Antecedents to the Adoption of Augmented Reality Smart Glasses: A closer Look at Privacy Risks", *Journal of Business Research* (92), pp. 374–384.

Rindova, V. P., and Petkova, A. P. 2007. "When Is a New Thing a Good Thing? Technological Change, Product Form Design, and Perceptions of Value for Product Innovations", *Organization Science* (18:2), pp. 217–232.

Rogers, E. M. 2003. *Diffusion of Innovations*, The Free Press, New York.

Tornatzky, L. G., and Klein, K. J. 1982. "Innovation Characteristics and Innovation Adoption-Implementation: A Meta-Analysis of Findings", *IEEE Transactions on engineering management* (1), pp. 28-45.

Venkatesh, V., Morris, M.G., Davis, G.B., and Davis, F.D. 2003. "User Acceptance of Information Technology: Toward a Unified View," *MIS Quarterly* (27:3), pp. 425-478.

Venkatesh, V., and Davis, F. D. 2000. "A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies," *Management Science* (46:2), pp. 186–204.

Yang, H., Yu, J., Zo, H., and Choi, M. 2016. "User Acceptance of Wearable Devices: An Extended Perspective of Perceived Value," *Telematics and Informatics* (33:2), pp. 256-269.