Consumers´ Acceptance of a Voice Commerce Application in FMCG in Germany, U.S. and U.K.

Elena Adolphs, Silvia Zaharia*

University of Applied Sciences Niederrhein, Krefeld, Germany
silvia.zaharia@hs-niederrhein.de

Abstract

Shopping-related voice assistant applications are on the rise, but their acceptance differs depending on the country. This paper examines customers' acceptance of a voice commerce application developed by a global fast moving consumer goods company based on a survey of online shoppers (n = 824) conducted in Germany, U.K. and the U.S. The main objective of the study is to identify which factors influence the acceptance of the voice commerce application, and whether there are differences between Germany, the U.S. and U.K. An integrated explanatory model was developed based on the Unified Theory of Acceptance and Use of Technology 2 (UTAUT2) with the antecedents: hedonic motivation, performance expectancy, effort expectancy and social influence. The original model was expanded to include the construct of perceived risk with the dimensions privacy and functional risk. The main result is that there are differences between the three countries regarding the factors that influence the acceptance of the voice commerce application. Only two factors have a significant influence in all three countries: performance expectancy and social influence, with performance expectancy demonstrating the strongest effect. From the perceived risks, only privacy risk has a negative influence on the intention to use the voice application in Germany. This study indicates that researches on the consumers' acceptance from one country should not be applied readily to another. It is rather advisable to consider the unique circumstances of each country.

Keywords: voice commerce · voice application · voice assistant · hedonic motivation · performance expectancy · effort expectancy · social influence · privacy risk · functional risk · intention to use · smart speaker

1 Introduction

Many consider voice commerce a revolution in online retailing. It enables new types of services to emerge and offers companies the opportunity to establish an exclusive and personal relationship with their customers [19]. Driven by mobile commerce, the usage of smart speakers, digital assistants and the ongoing implementation of the "Internet of things", this topic has become of significant relevance [11]. Voice commerce is a special form of e-commerce. It describes the interaction between users and commercial platforms and applications that utilize natural language speech recognition to enable self-service transactions over connected devices [18]. The devices used are

F. F.-H. Nah and K. Siau (Eds.): HCII 2021, LNCS 12783, pp. 3–21, 2021.
https://doi.org/10.1007/978-3-030-77750-0_1

equipped with conversational communication interfaces and intelligent software programs and are operated by the user using natural language [26]. The devices can be so-called smart speakers (such as Amazon Echo, Google Home or HomePod from Apple), but also computers or smartphones into which the software application of a voice assistant such as Alexa (Amazon), Google Assistant, Siri (Apple) or AliGenie (Alibaba) is integrated. Compared to previous voice-controlled human-computer interactions, the communication skills of artificial intelligence (AI) empowered voice assistants are far more advanced. Natural language processing enables people to talk to a computer / device like to a person and to receive contextual answers from them [6]. The two voice assistants covering commerce the most (in the western world) are Alexa and Google Assistant. Companies looking to sell their products using either of the voice assistants (VA) must develop a so-called shopping solution (for Alexa, these are called “skills” and for Google, “actions”).

For the consumer, the main benefits of using voice assistants are: convenience (they allow personalized, hands and eyes free usage [10] and are easy to use [12]), efficiency (less mental effort) and usage enjoyment [24]. The interaction and conversation with the artificial voice while shopping can be fun for the user. Over time, users can develop an emotional closeness to the devices and build a sense of a social relationship, similar to when interacting with people [7]. Consumers are also fascinated by the ability of the voice application to learn (due to the artificial intelligence of the voice assistant software) [9]. Alongside the benefits of using voice assistance, consumers hold several concerns regarding a lack of trust in the technology and a fear of losing privacy. The main barriers include: concerns about personal data (refusing to be actively recorded or personal data to be used) [10, 16], concerns about the functionality of the voice assistants (e.g. voice-only voice assistants don’t allow users to visualize information/choices or they do not understand the user and his/her reactions) [12]. Further concerns relate to the utility of voice assistants as well as the quality of the information provided [10].

Online retailers and consumer brands manufacturers expect the convenience and user experience offered by the voice assistants to have great potential in electronic commerce [19]. Over 60% of consumers in the U.S. who have an intelligent voice assistant have already used it for purchases. For the year 2022, revenue in the area of voice commerce for the U.S. market is estimated at over $ 40 billion [18]. Looking at the relevance of different branches in voice commerce, FMCG is a favorite product range for voice-assisted purchases [25].

Voice commerce applications can be used for more than simply purchasing products or services. Rather, they are relevant to all phases of the customer journey. Activities like “making a shopping list”, “researching a product/service”, “searching for a product/service", “comparing products/services" and “price comparison for products/services” are the most important reasons for using voice assistance in an e-commerce context [17]. However, this “market mediation“ function may also jeopardize brand manufacturers by changing their relationship to consumers. Voice assistants become “gate keepers”, whose built-in AI recommends specific brands to consumers based on their known preferences and purchasing behavior. Consumer brands fear a reduction in brand visibility via organic search results and in turn the rise of retailers’ private labels [13].

Despite the attention given to voice commerce from practitioners and industry reports, there is little academic research on the topic. There are some papers on the acceptance of voice assistants in general, but few empirical researches that relate to shopping (voice commerce). Most of these studies were merely carried out in one country, usually with a sample of students or university employees. The extent to which the results would differ between countries due to local peculiarities was not examined. Closing this research gap is the aim of this study.

The research object for this study is a German fast moving consumer goods (FMCG) manufacturer. The company in question is globally successful, with Germany, U.S. and the U.K. among its most important markets for beauty care products. The manufacturer pursues indirect distribution, does not have its own online shop and sells its products exclusively through offline and online retailers (e.g. Amazon). For this reason, the company does not aim to sell products via voice commerce. Instead, its goal is to develop a voice commerce application that calls consumers attention to its products early on in the customer journey to raise the visibility of the brand. For the purposes of this study, a prototype of a voice application was developed (use case) that dialogues with the user to suggest tailored beauty care solutions. With the help of Amazon skills or Google actions, the products can then be purchased through an associated online retailer. The present paper aims to answer the following research questions:
(1) Is there an intent to use the voice application?
(2) Which factors influence the intention to use the voice application and to what extent?
(3) Are there differences between the studied countries Germany, U.S. and U.K.?

2 Conceptual Framework

2.1 Technology Acceptance and Perceived Risk

Studies on the acceptance of voice assistants are based in part on the *Technology Acceptance Model* (TAM) by Davis et al. [4] as well as the *Unified Theory of Acceptance and Use of Technology* (UTAUT) and UTAUT2 by Venkatesh et al. [27, 28]. The UTAUT2 is the evolution of the UTAUT. It integrates eight established behavioral models of psychology and technology: The Theory of Reasoned Action, the Technology Acceptance Model, the Motivational Model, the Theory of Planned Behavior, the Combined TAM and TPB, the Model of PC Utilization, the Innovation Diffusion Theory and the Social Cognitive Theory [28]. The UTAUT2 represents an improved version of the UTAUT used to investigate use intentions (behavioral intention), thus better predicting the adoption (use behavior) of a technology [14]. The four determinants of UTAUT performance expectancy, effort expectancy, social influence and facilitating conditions were extended by hedonic motivation, price value and habit. The effect of performance expectancy on behavioral intention are moderated by the variables of age and gender. The effects of all other predictors (effort expectancy, social influence, facilitating conditions, hedonic motivation, price value and habit) on behavioral intention are moderated through the variables age, gender and experience [28].

When it comes to technology acceptance research, perceived risk has proven to be an important barrier. Featherman/Pavlou [5] used seven risk dimensions in their model and integrated the TAM in their study to find out how important perceived risk is for the decision to introduce e-services. The seven risk dimensions are: performance, financial, time, psychological, social, privacy and overall risk. Martins et al. [14] applied the model of Featherman/Pavlou [5] to the UTAUT of Venkatesh et al. [28] and developed a *Unified Theory of Acceptance and Use of Technology and Perceived Risk Application.*

2.2 Literature Review

Most studies on the acceptance / adoption of voice assistants do not relate explicitly to voice commerce [3, 7, 8, 10], or are of a qualitative nature [9, 16, 24]. Some current quantitative studies from the U.S., U.K. and Germany are presented below.

Liao et al. [10] conducted a survey with 1.178 users and non-users of voice-controlled intelligent personal assistants (IPAs) in the U.S. to investigate (1) the motivations and barriers to adopting IPAs and (2) how concerns about data privacy and trust in company compliance with social contract related to IPA data affect acceptance and use of IPAs. The adoption (or rejection) decisions are influenced by classical constructs in TAM and UTAUT: *perceived usefulness, performance expectancy* and *effort expectancy* associated with IPA use. Respondents who refused to consider purchasing a Home IPA (i.e. smart speaker) had significantly higher concerns about the use of data and a significantly lower confidence that the data is sufficiently secure.

McLean/Osei-Frimpong [15] take a Uses and Gratification Theory (U&GT) approach to explain the use of in-home voice assistant with a sample of 724 users in the U.K. The results from a structural equation model illustrate that individuals are motivated by the *utilitarian benefits*, *symbolic benefits* and *social benefits* provided by voice assistants. The *hedonic benefits* do not motivate the use of in-home voice assistants. Additionally, the research establishes a moderating role of *perceived privacy risks* in dampening and negatively influencing the use of in-home voice assistants.

Wagner et al. [30] investigated (1) the role of anthropomorphism in the context of digital voice assistants and (2) the determinants of the UTAUT2 for digital voice assistants by conducting an online survey with 283 users. The results of the structural equation modelling show (1) that anthropomorphism in general plays a role concerning the behavioral intention for voice assistants with a *humanlike-fit* having the highest impact on a human driven likeability. (2) The relevant drivers of the intention to use voice assistants (referring to the UTAUT2 model) are *performance expectancy*, *hedonic motivation* and *habit. Facilitating conditions, effort expectancy* and *social influence* had no significant influence on the intention to use voice assistants.

Two relevant <u>voice commerce</u> studies based on UTAUT2 were conducted at the University of Applied Sciences Niederrhein in Germany. Puschmann et al. [22] used multiple regression analysis to study the acceptance of voice assistants in e-commerce (n = 429). The key findings of the analysis are that *performance expectancy* and *hedonic motivation* have the strongest influence on the intention to use voice assistants for purchases (*behavioral intention*). Meanwhile, *effort expectancy* and *social influence* only have a low influence on purchasing intentions. *Facilitating conditions* had no significant influence on the intention to use, while *perceived risk* had a weak negative effect on it. The influence of the moderators (age, gender and experience) on the relationship between predictors and behavioral intention proved significant for only two predictors (performance expectancy and hedonic motivation): The older a person is, the weaker the influence of performance expectancy on behavioral intention. Meanwhile, the effect of performance expectancy and hedonic motivation on behavioral intention is weaker for women than for men.

By means of a representative online survey of German online shoppers (n = 684), Zaharia/Würfel [33] studied the factors influencing the acceptance of smart speakers throughout the customer journey. The explanatory model developed for the study was based on UTAUT2 and expanded to include the construct of perceived risk. The examined structural equation model revealed *performance expectancy* and *hedonic motivation* as the strongest factors influencing the willingness of online shoppers to use smart speakers (*behavioral intention)*. Prior *experience* and the perceived *price value* of smart speakers had little effect on the intention to use them in voice commerce. *Effort expectancy* had no direct effect on behavioral intention, while *perceived risk* had a negative effect on the intention to use. Furthermore, the intention to use smart speakers was shown to be higher during the information phase than in the purchasing phase.

2.1 Model and Hypotheses

The present study builds on the results of Zaharia/Würfel [33] and Puschmann et al. [22]. The basic model is the UTAUT2, extended to include the construct perceived risk. Ideally, the acceptance of voice commerce applications would be measured on the basis of actual use behavior. However, the application in question exists only as a prototype shown to respondents for the first time in the context of the study. Instead, the dependent variable for measuring acceptance is *behavioral intention* (BI). Users are meant to utilize the selected voice commerce application during the pre-sales phase.

In developing the model for this study, three independent variables were eliminated from the UTAUT2 model (facilitating conditions, price value and habit): According to Puschmann et al. [22] and Wagner et al [30], the variable facilitating conditions is a not significant factor. Price value as an independent variable is not relevant to this study because the software is free of charge, and a consumer need not buy a smart speaker to use the voice commerce application. Furthermore, the actual voice application (the use case) cannot be related to habit, as it was not used by the consumer prior to the study. The model developed can be seen in Figure 1. *Perceived risk* (*functional* and *privacy*) is assumed to have a negative effect on the intention to use the presented voice application (*behavioral intention*). The factors *performance expectancy* (PE), *effort expectancy* (EE), *social influence* (SI) and *hedonic motivation* (HM) have a positive influence on the *behavioral intention* (BI). *Age* and *gender* moderate the influence of PE, EE, SI and HM on BI.

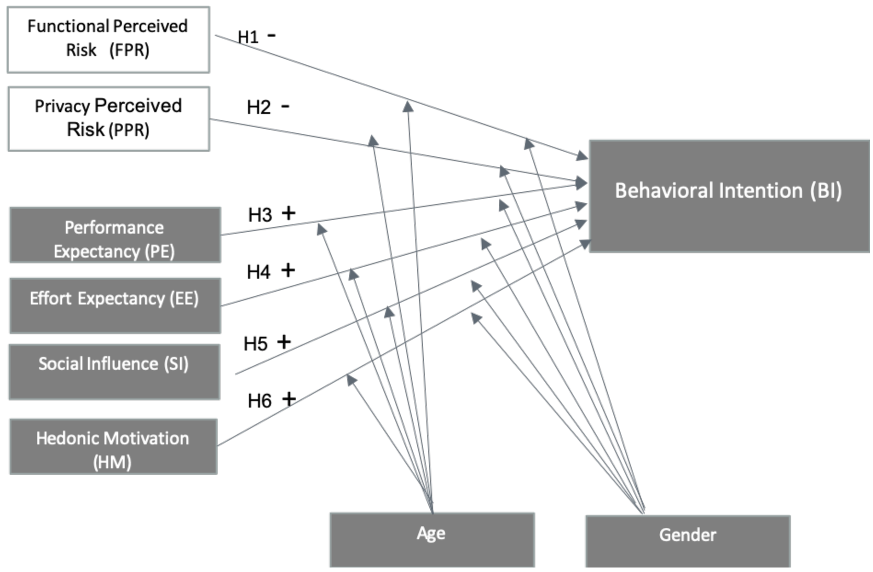


Fig. 1: Explanatory model for the acceptance of a voice commerce application

Perceived risk is defined as the degree to which a user thinks that using a technology will have negative implications for him or her and has a negative impact on the intention to use a technology [14]. Consumers have been shown to hold several concerns about the usage of voice assistants. Liao et al. [10] found that perceptions of whether IPA providers adhere to privacy and security rules affects users' likelihood of using IPAs, even if only 7% of respondents cited privacy concerns as the main reason for not using IPAs. An exploratory German study names the following negative beliefs toward voice commerce [24]:

- limited transparency (no visual representation, no comparison function, limited product information)
- low technical maturity (limited interactivity, speech recognition errors)
- limited control (potential misuse by strangers, no manual input modality, risk of misunderstanding)
- lack of trust (vendor's competence / benevolence, technology reliability).

The different facets of perceived risk can be summarized by the two dimensions of *functional perceived risk (FPR)* and *privacy perceived risk* (PPR). For the moderators age and gender, it is found that the major demographic using voice assistants is 33-45 years old [2] and male [23]. Ipso facto, it is assumed that this group perceives less functional and privacy risk than older people and women. This leads to the following hypotheses:

H1: The influence of FPR on BI is negative.
H1a: The influence of FPR on BI is weaker for younger people.
H1b: The influence of FPR on BI is stronger for women.
H2: The influence of PPR on BI is negative.
H2a: The influence of PPR on BI is weaker for younger people.
H2b: The influence of PPR on BI is stronger for women.

Performance expectancy (PE) means the degree to which a consumer expects to experience a performance advantage (utility) from using a voice commerce application. Consumers expect an advantage in terms of convenience and time savings by using voice assistants [33]. Furthermore, the available functions of the voice assistants have a positive influence on the adoption of the technology [9]. This leads to the assumption that a positive PE increases the usage intention of the use case. For the moderators age and gender, studies demonstrate that women are less task-oriented than men. This suggests that it is less important for them to implement a task in a targeted way. For age it is mentioned that younger people tend to attach greater importance to extrinsic rewards [28]. This leads to the following hypotheses:
H3: The influence of PE on BI is positive.
H3a: The influence of PE on BI is stronger for younger people.
H3b: The influence of PE on BI is weaker for women than men.

Effort Expectancy (EE) refers to the amount of effort expected when using a technology. In the voice commerce context, it is defined as the degree to which a consumer considers an application easy to learn and operate [29, 33]. According to UTAUT2 the assumption is that a high usability of the voice commerce application will have a positive influence on BI. For the moderator gender it was found that there is a gender difference and that ease of learning a technology is more important for women than for men [31]. For age it can be detained that younger people have a higher cognitive capacity for innovation [9]. This leads to the following hypotheses:
H4: The influence of EE on BI is positive.
H4a: The influence of EE on BI is stronger for younger people.
H4b: The influence of EE on BI is stronger for women.

Social influence (SI) means the degree to which a consumer experiences important people (family and friends) recommending the voice assistant application [22]. Yang et al [32] explain that four social influenced indicators affect the individual's intention to use a technology: subjective norm, image, visibility and voluntariness. The presented use case cannot refer to image, visibility or voluntariness effects, but it can refer to subjective norms. Subjective norms means that individuals are influenced by other people and prefer them as a source for information and guidance [32]. People mention an increased interest in voice assistants, if a friend has told them about it [9]. This leads to the assumption that BI towards the use case is increased by SI. This effect is moderated by age and gender because women and older people tend to attach greater importance to the opinions of others [27]. This leads to the following hypotheses:

H5: The influence of SI on BI is positive.
H5a: The influence of SI on BI is stronger for older people.
H5b: The influence of SI on BI is stronger for women.

Hedonic motivation (HM) refers to the perceived pleasure an individual experience from using a technology [9]. In the voice commerce context, HM is defined as the degree to which a consumer considers using voice commerce applications as fun, entertaining, exciting and pleasant [33]. According to previous research, it is assumed that a high degree of fun and satisfaction while using the voice application will lead to a higher intention to use it (BI). Age and gender moderate the fun and enjoyment arising from new technology use. It tends to be stronger for men, as men have a greater interest in innovative systems and are more open-minded to integrating them [1]. This is emphasized by studies showing that voice assistants are used by a majority of men [23]. Furthermore, younger people are more curious about discovering new technologies and older people are more skeptical of technological innovation [21]. This leads to the following hypotheses:
H6: The influence of HM on BI is positive.
H6a: The influence of HM on BI is stronger for younger people.
H6b: The influence of HM on BI is weaker for women.

3 Research Design

3.1 Presenting the Use Case

In order to develop a voice commerce application for this study, a qualitative survey was conducted as a first step. Researchers studied voice commerce applications already on the market (from competitors, retailers or best practices from other industries) and interviewed experts. This resulted in a mockup voice application (use case) that recommends tailored product solutions by dialoging with users. The voice application is a hair advisor based on visual content that targets consumers looking for new hair styles and hair trends. Users will use it in the bathroom or in favorite spaces like the living room or kitchen on display devices like smartphones, Google Nest Hub and Amazon Echo Show. The skill was launched for both the voice assistants Alexa and Google Assistant. The user contacts the skill or action by name and call to action i.e. “Alexa/Google open the Hair Application”. The skill can also be found by asking “Alexa/Google show me the newest hairstyle”, “Alexa/Google, I want to color my hair”, “Alexa/Google how can I braid/curl/straighten my hair?”. The aim is to attract customers by identifying their preferred looks before they begin searching for products. The products linked to each look (e.g. hair color) are then suggested to the user. The user can look up product information for each product and put items on their shopping list if using the Amazon voice assistant. In Google’s case the phrase “put it on a shopping list” will advise the user on where to purchase the product by showing and redirecting the customer to the top three online retail shops where the product is available.

3.2 Data Collection

The quantitative data of the research was obtained from a sample of 824 participants in an online survey conducted in September 2019. The studied countries were Germany (DE; n = 281), United Kingdom (U.K.; n = 286) and the United States (U.S.; n = 257). The sample represents the FMCG company's target group: men and women from each country ages 16-65 years who are open to new technologies (e.g. voice commerce/voice assistants) and who have a basic interest in hair styling. Quotas were set in order to reach a sufficient case number within each gender and age group. For age it was equally 33% through the main age groups 16-29, 30-49 and 50-65 years. For gender the quotas were 90% women and 10% men to cover the target group.

The survey starts with presenting the use case (mockup) to the respondent followed by questions on the actual use of voice assistance and voice skills. Further questions target each individual construct. The constructs were measured using multi-item scales on a 5-point Likert Scale (1 = "I strongly disagree" / 5 = "I strongly agree"). The operationalization of the constructs can be seen in Appendix 1.

4 Results

The descriptive evaluation answered the first research question "Is there an intent to use the voice application?". The top-two analysis shows that, in total, 54% of consumers agree with the statement "I intend to use the application in the future". However, responses differ at a country level. The highest share of people who state intentions to use the application are in the U.S. (70%) followed by U.K. (51%). Germans seem to be the most hesitant with only 41%. To learn more about the general usage, two additional questions were stated in the survey asking the respondents for the actual use of voice assistance and voice skills. Most respondents do not yet use voice assistance (36 %) nor voice skills (40%).

To verify the constructs, a confirmatory factor analysis (CFA) was performed on the total database. A data quality check and testing of univariate exploratory factor analysis (EFA) was done in advance [34]. The total result can be summarized as very good, having used basically proven operational questions from existing models. The following test results are valid (see table 1):

1. Cronbach's Alpha refers to the joint consideration of the Item-to-Total-Correlation (ITC). Depending on the number of items α limits are between ≥0.5 -≥0.7 (2 items $\alpha \geq 0.5$; 3 items $\alpha \geq 0.6$; more than 4 items $\alpha \geq 0.7$.) The limit values are maintained for all constructs.
2. By using exploratory factor analysis, the constructs are tested for reliability and validity. The factor loadings should show a correlation between item and factor of at least 0.7. The indicator reliability should be above the value of 0.5 in the form of communalities. The test by univariate EFA is positive for all items.
3. By performing CFA (AMOS), the significance of the factor loads and the standardized factor loads are determined. From this the quality criteria of the average

variance extracted (AVE ≥ 0.5) and the factor reliability (FR ≥ 0.6) are calculated and tested. All factor loads could be tested as significant.

4. The Fornell-Larcker criterion determines the discriminant validity of the construct to the other constructs, thus demonstrating the quality of the factor. To substantiate this, the average variance of a factor must be higher than the factors squared correlation and other measured factors. This can be supported for all factors used in the model.

The hypotheses were tested using multiple regression analysis. Multiple regression is linked to a number of assumptions about the nature of the data [34]. These assumptions were fulfilled for all three countries. Testing the model revealed that the variance of BI (the probability to use the voice application) can be explained by 67.8% (DE), 74.4% (U.K.) and 65.8% (US) of the variance of the predictors. At a 0.01% level, the model contributes to explaining the regressor for all countries (ANOVA).

The results of the regression analysis (see table 2) answer the second research question: "Which factors influence the intended use?" Appendix 2 shows the effect model for each country. Germany is the only country showing an effect (negative) from perceived risk. In all three countries, the effects of the predictors are weak to moderate.

- For Germany, *the intention to use the voice application* is significantly influenced by *performance expectancy*, *social influence* and *hedonic motivation*. Meanwhile, *perceived privacy risk* has a weak negative effect.
- In the U.K. only *performance expectancy*, *social influence* and *hedonic motivation* have a significant effect on the *intention to use the voice application*.
- The results from the U.S. differ. *Performance expectancy*, *effort expectancy* and *social influence* have a significant effect on the *intention to use the voice application.*

The analysis shows that all main hypotheses, with the exception of H1, could be supported for nearly all cases across the three countries (see Table 3).

Table 1 Confirmatory factor analysis results including quality criteria

Construct	α ≥ 0.5-0.7	Average variance (AVE) ≥ 0.5	FR ≥ 0.6	1-factorial solution	Explained variance in %	Indicator	ITC ≥ 0.3-0.5	α if Item is deleted	Sign. Factor loading	Factor loading ≥ 0.7	Commonalities ≥ 0.5
Functional Perceived Risk	0.912	0.812	0.945	yes	79.124	FPR_1	0.785	0.891	***	0.880	0.775
						FPR_2	0.781	0.892	***	0.878	0.770
						FPR_3	0.823	0.878	***	0.904	0.817
						FPR_4	0.810	0.883	***	0.896	0.803
Privacy Perceived Risk	0.796	0.657	0.847	yes	71.009	PPR_1	0.556	0.804	***	0.784	0.614
						PPR_2	0.703	0.65	***	0.880	0.775
						PPR_3	0.669	0.69	***	0.861	0.741
Performance Expectancy	0.875	0.784	0.916	yes	80.199	PE_1	0.787	0.799	***	0.909	0.827
						PE_2	0.752	0.834	***	0.890	0.792
						PE_3	0.748	0.835	***	0.887	0.787
Hedonic Motivation	0.875	0.797	0.921	yes	80.135	HM_1	0.807	0.78	***	0.922	0.849
						HM_2	0.813	0.774	***	0.925	0.855
						HM_3	0.665	0.906	***	0.836	0.699
Effort Expectancy	0.900	0.780	0.934	yes	77.078	EE_1	0.714	0.894	***	0.833	0.694
						EE_2	0.811	0.859	***	0.899	0.809
						EE_3	0.809	0.86	***	0.898	0.807
						EE_4	0.777	0.871	***	0.879	0.773
Social Influence	0.868	0.745	0.921	yes	71.634	SI_1	0.651	0.857	***	0.799	0.638
						SI_2	0.733	0.826	***	0.855	0.731
						SI_3	0.756	0.815	***	0.869	0.756
						SI_4	0.739	0.823	***	0.860	0.740
Behavioral Intention	0.942	0.899	0.964	yes	89.588	BI_1	0.877	0.917	***	0.946	0.895
						BI_2	0.846	0.941	***	0.930	0.864

(ns) = not significant/ *significance on 5% level/ ** high significance on 1% level/ *** highly significance on 0.1% level

F. F.-H. Nah and K. Siau (Eds.): HCII 2021, LNCS 12783, pp. 3–21, 2021. https://doi.org/10.1007/978-3-030-77750-0_1

Table 2 Results of the regression analysis

Costruct	Dimension	B			(Beta)			Significance						adjR²		
		DE	*U.S.*	*U.K.*	*DE*	*US*	*U.K.*	*DE*		*U.S.*		*U.K.*		*DE*	*U.S.*	*U.K.*
Behavioral Intention	(Constant)	0.060	-0.377	-0.497				0.816	(ns)	0.122	(ns)	0.043	(ns)	0.678	0.658	0.744
	(FPR) Functional Perceived Risk	0.121	0.030	0.053	0.097	0.039	0.044	0.016	*	0.379	(ns)	0.271	(ns)			
	(PPR) Privacy Perceived Risk	-0.150	0.004	-0.094	-0.139	0.005	-0.072	0.001	***	0.908	(ns)	0.068	(ns)			
	(PE) Performance Expectancy	0.423	0.540	0.464	0.391	0.487	0.400	0.000	***	0.000	***	0.000	***			
	(HM) Hedonic Motivation	0.208	0.009	0.252	0.187	0.007	0.206	0.007	**	0.919	(ns)	0.001	*			
	(EE) Effort Expectancy	0.004	0.312	0.031	0.003	0.252	0.022	0.950	(ns)	0.000	**	0.685	(ns)			
	(SI) Social Influence	0.390	0.203	0.361	0.348	0.241	0.345	0.000	***	0.000	***	0.000	***			

(ns) = not significant/ *significance on 5%-level/ ** high significance on 1% level/ *** highly significance on 0.1% level

medium effect (> 0.3) weak effect (> 0.1)

Table 3 Test result of the main hypotheses

Hypothesis		Assessment *DE*	*U.S.*	*U.K.*
H1	The influence of FPR on BI is negative.	x	x	x
H2	The influence of PPR on BI is negative.	**supported**	x	x
H3	The influence of PE on BI is positive.	**supported**	**supported**	**supported**
H4	The influence of EE on BI is positive.	x	**supported**	x
H5	The influence of SI on BI is positive.	**supported**	**supported**	**supported**
H6	The influence of HM on BI is positive.	**supported**	x	**supported**

x = not supported

The effect of the moderators age and gender on respective predictor-criterion relationships was examined using multiple group regression analysis. For gender, the moderating effect was not assessed by country since, after data cleaning, fewer than 30 men per country remained (D: n = 24 and U.K.: n = 28). In reviewing the requirements for the regression analysis, a problem arose in the VIF value for HM (5.571), meaning the moderator effect of age and gender for the influence of HM on BI could not be assessed. The hypotheses involving moderating effects could not be supported in most cases (see Appendix 3).

The third research question "Are there differences between the studied countries Germany, U.S. and U.K.?", can be answered affirmatively. Differences arise both in the use intention as well as the triggers and barriers for the intention to use the voice commerce application.

5 Discussion, Recommendations and Limitations

The present study has shown that there are differences between Germany, the U.S. and the U.K. when it comes to acceptance of the voice commerce application. The highest intention to use was seen in the U.S. (70%) followed by the U.K. (51%). In Germany, only 41% stated an intention to use the voice commerce application. From the UTAUT2 antecedents only two were shown to have an influence in all three countries, namely performance expectancy and social influence.

- Performance expectancy has the strongest effect in all countries and was shown to be higher for men.
- The effect of social influence on the intention to use the voice commerce application is higher in DE and the U.K. than in the U.S. For age differentiation, the influence of social influence was shown to be stronger for older people in Germany. The social influence effect is in all three countries higher for women.
- Effort expectancy has an influence only in the U.S. where it demonstrated the second strongest effect.
- Hedonic motivation only has an influence in Germany and the U.K.
- Privacy risk has a negative influence only in Germany. For all other countries, the risk predictors are not significant.

This raises the question if, and to what degree, studies from one country can be applied readily to another. This study shows that it is advisable to consider the specific circumstances of each country when developing and implementing the application. Since users´ performance expectancy (i.e. the perceived benefit to him or her from using the voice commerce application) has the strongest effect on acceptance in all countries, success hinges on the ability to ensure performance of the voice commerce application. The user will rate the voice application primarily according to how well it helps him/her to find the right product. This means that the intelligent software should make the right product suggestions based on the input from the consumer and provide the desired information.

Social influence is in all three countries very important for implementing the application, especially for women. This key trigger should be pushed in all three countries. This can be done by using a recommendation function in the app. The intention to use the application can for example be influenced by sweepstakes on social media where people are asked to test the application and tell about it.

Due to the differences between the three countries, the following country-specific measures are recommended: In the U.S., the effortlessness (easy to understand and operate) of the application should be underlined. For the U.K. and Germany, where hedonic motivation plays a role in acceptance, communication should focus on the fun experienced while using the application. In order to guarantee the perceived enjoyment of use and to avoid frustration, companies should ensure the integration of the voice channel with internal processes. In Germany in particular, the perceived privacy risk must be taken into account. Companies should be proactive to build trust and ensure and communicate the privacy and security of customer data.

There is a number of limitations in our research study that should be addressed in future research. The first limitation pertains to the research object: the results of the study are relevant for FMCG manufacturers in the area of beauty care products and their voice commerce application. Results for a different voice commerce application in another industry could turn out quite differently. The second limitation pertains to the countries included in the study. To generate the most relevant findings, the study focused on the three selected countries because they are the most important for e-commerce for the FMCG manufacturer. Still, other relevant markets should be considered for future study. China is an interesting option since it is indicated as open minded for new technologies.

6 Conclusion

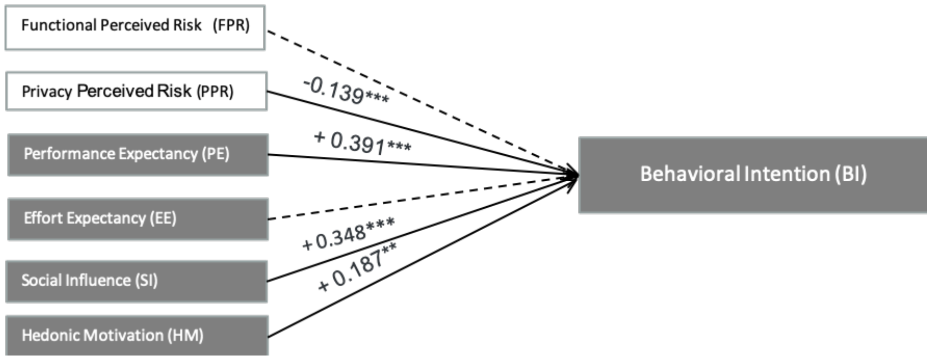
Voice commerce applications carry opportunities and risks for consumers, but they are weighed differently depending on the country. This study examined customers' acceptance of a voice commerce application developed by a global FMCG Company based on an online survey of online shoppers (n = 824) conducted in Germany, U.K. and the U.S. An integrated explanatory model was developed based on the UTAUT2. The original model was expanded to include the construct of perceived risk with the dimensions privacy and functional risk. The proposed conceptualization and operationalization of the constructs was analyzed by means of exploratory and confirmatory factor analysis and found to be very good (based on generally recognized quality criteria). The hypothesized relationships were examined with the help of regression analysis. The present study has shown that there are differences between the three countries regarding the factors that influence the intention to use the voice commerce application. Performance expectancy and social influence have a significant influence in all three countries, with performance expectancy demonstrating the strongest effect in all countries. The influence of performance expectancy is higher for men while social influence is higher for women. From the perceived risks, only privacy risk has a weak but highly significant negative influence on the intent to use the voice application in Germany. Perceived functional risk has no significant effect. Effort expectancy has an influence only in the U.S. where it has the second strongest effect. Hedonic motivation only has an influence in Germany and the U.K. This study shows that it is advisable to consider the specific circumstances of each country when developing and implementing voice commerce applications.

Appendix 1: Operationalization of the constructs

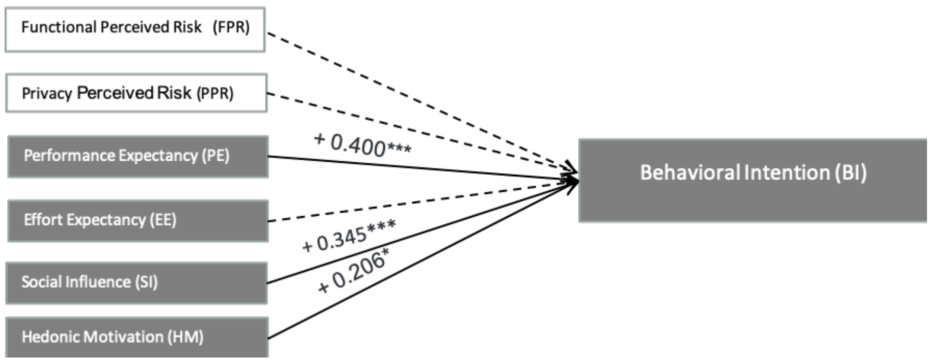
Functional Perceived Risk	
FPR_1	*I suspect the voice application (VA) could…* …not perform well and create problems with my devices.
FPR_2	…highly probably not provide desired results.
FPR_3	…advise me wrong, because of malfunction.
FPR_4	…advise me wrong, because the assistant will not understand me.
Privacy Perceived Risk	
PPR_1	Using the VA with my voice assistant will cause my conversations to be over-heard.
PPR_2	Signing up for and using the application would lead to a loss of privacy because my personal information would be used without my knowledge.
PPR_3	Internet hackers (criminals) might take control of my checking account if I use the VA.
Performance Expectancy	
PE_1	*When I think of the VA, I would assume…* …that I would find it useful in my daily life.
PE_2	...it to enable me to answer my questions about hair styling and hair products more quickly.
PE_3	...it to increase my productivity because I can do several things at once.
Effort Expectancy	
EE_1	The interaction with the application is clear and understandable.
EE_2	It is easy for me to become skillful at using the application.
EE_3	I find the application easy to use.
EE_4	Learning how to use the application would be easy for me.
Social Influence	
SI_1	*Whether I will use the VA in the future could be influenced by…* …friends or family members recommending it to me.
SI_2	…Influencers recommending it to me via social media (e.g. Facebook or Insta-gram).
SI_3	…very important people recommending it to me via advertising.
SI_4	... colleagues and superiors whose opinion I value recommending it to me.
Hedonic Motivation	
HM_1	Using the application is fun.
HM_2	Using the application is entertaining.
HM_3	The VA supports me in the shopping process (e.g. suggest products to me and put them in the shopping cart).
Behavioral Intention	
BI_1	I intend to use the application in the future.
BI_2	I will always try to use the application in my daily life.
BI_3	I plan to use the application frequently.

Appendix 2: The effect model for each country

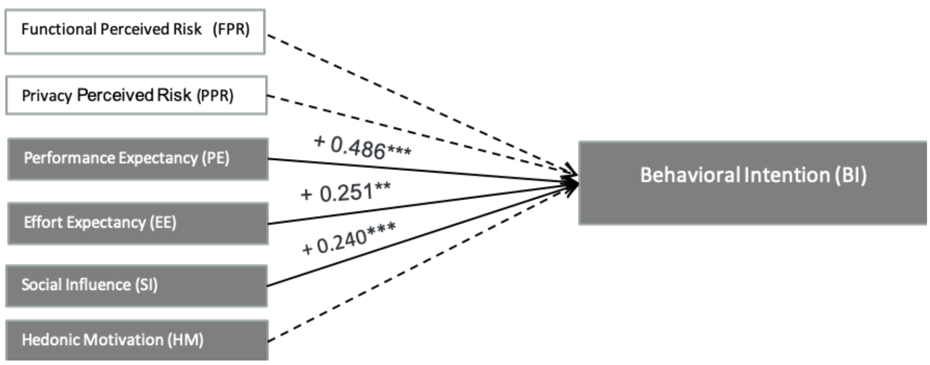
Germany



U.K.



U.S.



Appendix 3: Test result of hypotheses for moderating effects

Hypothesis	DE	U.S.	U.K.
H1a: The influence of FPR on BI is weaker for younger people.	x	x	x
H1b: The influence of FPR on BI is stronger for women.		x	
H2a: The influence of PPR on BI is weaker for younger people.	x	x	x
H2b: The influence of PPR on BI is stronger for women.		x	
H3a: The influence from PE on BI is stronger for younger people.	x	x	x
H3b: The influence of PE on BI is weaker for women than men.		supported	
H4a: The influence of EE on BI stronger for younger people.	x	x	x
H4b: The influence of EE on BI is stronger for women.		x	
H5a: The influence of SI on BI is stronger for older people.	supported	x	x
H5b: The influence of SI on BI is stronger for women.		supported	

x = not supported

7 References

1. Burke, R. R. (2002). Technology and the customer interface: what consumers want in the physical and virtual store. Journal of the Academy of Marketing Science, 30(4), 411–432.
2. Capgemini (2018). Conversational Commerce: Why Consumers Are Embracing Voice Assistants. https://www.capgemini.com/resources/conversational-commerce-dti-report/
3. Coskun-Setirek, A. & Mardikyan, S. (2017). Understanding the Adoption of Voice Activated Personal Assistants; in: International Journal of E-Services and Mobile Applications; 9/3
4. Davis, F. D., Bagozzi, R. P. & Warshaw, P. R. (1989): User Acceptance of Computer Technology: A Comparison of Two Theoretical Models; in: Management Science; 35 (1989); 8; S. 982–1003.
5. Featherman, M. S., & Pavlou, P. A. (2003). Predicting e-services adoption: a perceived risk facets perspective. International Journal of Human-Computer Studies, 59(4), 451–474.
6. Guzman, A. L. (2019). Voices in and of the machine: Source orientation toward mobile virtual assistants, in: Computers in Human Behavior, Volume 90, January 2019, Pages 343-350
7. Han, S. & Yang, H. (2018): Understanding adoption of intelligent personal assistants; in: Industrial Management & Data Systems; 118 (2018); 3; S. 618–636.
8. Kääriä, A. (2017). Technology acceptance of voice assistants: Anthropomorphism as a factor – Master thesis; Jyväskylä: University of Jyväskylä.
9. Kessler, S. K., & Martin, M. (2017). How do potential users perceive the adoption of new technologies within the field of Artificial Intelligence and Internet-of-

Things? - A revision of the UTAUT2 model using Voice Assistants - Master thesis Lund University, Sweden.
10. Liao, Y.; Vitak, J., Kumar, P., Zimmer, M. & Kritikos, K. (2019). Understanding the Role of Privacy and Trust in Intelligent Personal Assistant Adoption, in: Proceedings of the 14th International Conference, iConference 2019, Washington, DC, USA, March 31–April 3.
11. Loweg, R. (2019). Sprachassistenten vermehren sich rasant. https://www.wiso-net.de/document/CID__155452.
12. Luger, E., & Sellen, A. (2016). Like Having a Really Bad PA: the Gulf Between User Expectation and Experience of Conversational Agents, in: Proceedings of the CHI Conference on Human Factors in Computing Syst., pp. 5286-5297.
13. Mari A., Mandelli A., Algesheimer R. (2020). The Evolution of Marketing in the Context of Voice Commerce: A Managerial Perspective. In: Nah FH., Siau K. (eds) HCI in Business, Government and Organizations. HCII 2020. Lecture Notes in Computer Science, vol 12204. Springer, Cham. https://doi.org/10.1007/978-3-030-50341-3_32.
14. Martins, C., Oliveira, T., & Popovič, A. (2014). Understanding the Internet banking adoption: A unified theory of acceptance and use of technology and perceived risk application. International Journal of Information Management, 34(1), 1–13.
15. McLean, G., Osei-Frimpong, K. (2019). Hey Alexa … examine the variables influencing the use of artificial intelligent in-home voice assistants, in: Computers in Human Behavior, Volume 99, October 2019, Pages 28-37
16. Moorthy, A. E./Vu, Kim-Phuong L. (2015): Privacy Concerns for Use of Voice Activated Personal Assistant in the Public Space; in: International Journal of Human- Computer Interaction; 31 (2015); 4; S. 307–335.
17. Olson, C., & Kemery, K. (2019). Voice report: From answers to action: customer adoption of voice technology and digital assistants. https://printkr.hs-niederrhein.de:2069/study/63323/digital-assistant-and-voice-assistant-adoption-2019
18. Paluch S., Wittkop T. (2020). Voice Marketing – Die Stimme der Zukunft?, in: Bruhn M., Burmann C., Kirchgeorg M. (eds) Marketing Weiterdenken. Springer Gabler, Wiesbaden. https://doi.org/10.1007/978-3-658-31563-4_26
19. Piacenza, B., Kress, R., & Klämt, M. (2018). Voice Search Engine Optimization: Guidebook.
20. Pohlgeers, M. (2019). Voice-Commerce ist die größte Revolution im Online-Handel seit dem Smartphone. https://www.onlinehaendler-news.de/digital-tech/innovationen/131110-voice-commerce-groesste-revolution-online-handel-seit-smartphone?utm_source=newsletter&utm_medium=email&utm_campaign=expertsnl%20(accessed%2005/06/2019).
21. Prein, J. (2011). Akzeptanz mobiler Kundenkartenprogramme bei Konsumenten. Diss. Univ. Kassel. Wiesbaden: Gabler Verlag.
22. Puschmann, T., Bank, K. von der, Winterwerber, T., Dürbaum, A., & Birven, L. (2019). Akzeptanz von Voice Technologien im E-Commerce. Krefeld. Hochschule Niederrhein.
23. Riaz, N. (2018). Im Redeschwall. https://www.lead-digital.de/im-redeschwall/
24. Rzepka, C., Berger, B., Hess, T. (2020). Why Another Customer Channel? Consumers' Perceived Benefits and Costs of Voice Commerce, in: Proceedings of the

53rd Hawaii International Conference on System Sciences (HICSS 2020), Wailea, Hawaii, USA, January 7-10.
25. Statista (2019). Voice commerce in the United States. https://www.statista.com/study/60607/voice-commerce-in-the-united-states/
26. Tuzovic, S./Paluch, S. (2018): Conversational Commerce – A New Era for Service Business Development?; in: Bruhn, M./Hadwich, K. (ed.): Service Business Development – Strategien - Innovationen - Geschäftsmodelle : Band 1; Wiesbaden, Germany: Springer Gabler; S. 81–100.
27. Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS Quarterly, 425–478.
28. Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Quarterly, 157–178.
29. Wagner, M. (2016). Entwicklung und Überprüfung eines konsolidierten Akzeptanzmodells für Lernmanagementsysteme lmu]. EndNote Tagged Import Format.
30. Wagner, K., Nimmermann, F., Schramm-Klein, H. (2019). Is It Human? The Role of Anthropomorphism as a Driver for the Successful Acceptance of Digital Voice Assistants, in: Proceedings of the 52rd Hawaii International Conference on System Sciences.
31. Wang, H.-Y.,& Wang, S.-H. (2010). User acceptance of mobile internet based on the unified theory of acceptance and use of technology: Investigating the determinants and gender differences. Social Behavior and Personality: An International Journal, 38(3), 415–426.
32. Yang, H.-D., Moon, Y. J., & Rowley, C. (2009). Social influence on knowledge worker's adoption of innovative information technology. Journal of Computer Information Systems, 50(1), 25–36.
33. Zaharia, S.; Würfel, M. (2020). Voice Commerce - Studying the Acceptance of Smart Speakers, Proceedings of the IHIET 2020, Springer, https://doi.org/10.1007/978-3-030-55307-4_68.
34. Zinnbauer, M., & Eberl, M. (2004). Die Überprüfung von Spezifikation und Güte von Strukturgleichungsmodellen: Verfahren und Anwendung, München.