Routledge
Taylor & Francis Group

Communication Research Reports

ISSN: 0882-4096 (Print) 1746-4099 (Online) Journal homepage: www.tandfonline.com/journals/rcrr20

Hey Alexa, why do we use voice assistants? The driving factors of voice assistant technology use

Emily Buteau & Joonghwa Lee

To cite this article: Emily Buteau & Joonghwa Lee (2021) Hey Alexa, why do we use voice assistants? The driving factors of voice assistant technology use, Communication Research Reports, 38:5, 336-345, DOI: 10.1080/08824096.2021.1980380

To link to this article: https://doi.org/10.1080/08824096.2021.1980380

COMMUNICATION RESEARCH REPORTS
2021, VOL. 38, NO. 5, 336–345
https://doi.org/10.1080/08824096.2021.1980380

Routledge
Taylor & Francis Group

Hey Alexa, why do we use voice assistants? The driving factors of voice assistant technology use

Emily Buteau and Joonghwa Lee

ABSTRACT

This study identified three factors to predict the use of the artificial intelligence communication devices known as voice assistants such as Amazon's Alexa or Apple's Siri: technological factors (usefulness and perceived ease of use), social influence factors (personal and societal norms), and risk factors (privacy concerns and perceived security). Considering these factors as significant antecedents, this study proposed an extended Technology Acceptance Model. To examine the proposed model, an online survey was conducted with 558 respondents. The findings indicate that perceived usefulness, personal norms, and perceived security have positive relationships with attitudes toward using voice assistants. Additionally, privacy concerns had a negative relationship with attitudes toward using voice assistants, which in turn has a positive relationship with the behavioral intention. Practical and theoretical implications are discussed.

KEYWORDS

Voice assistants; technology acceptance model; personal injunctive norms; personal descriptive norms; privacy concerns

Voice assistants are artificial-intelligence devices that listen for commands and engage in conversational-based communication to assist and complete tasks for users (Hoy, 2018). Voice assistants can be accessed through devices like cell phones (i.e., Apple Siri), or can be purchased as a separate piece of technology (i.e., Amazon Alexa, Google Home) (Pitardi & Marriott, 2021). Common voice assistant tasks include playing music, completing purchases, sending messages, and more (Guzman, 2019). Although voice assistants are convenient and popular, there has been resistance due to security and privacy concerns such as collection of personal information (NPR, Edison Research Center, 2020), the use and storage of personal data (Vimalkumar, Sharma, Singh, & Dwivedi, 2021), and others listening while using the device (Moorthy & Vu, 2014).

Despite the increasing conversation regarding voices assistants, there are limited studies exploring the specific factors, aside from privacy concerns, influencing people to use them. Thus, the objective of this study is to identify the several factors leading to voice assistant acceptance as an extension to Technology Acceptance Model (TAM) (Davis, 1989). This study conceptualizes technological, social influence, and risk factors that predict the use of voice assistants as an extension of TAM. It is vital to understand voice assistant

CONTACT Joonghwa Lee joonghwa.lee@und.edu Department of Communication, University of North Dakota, Columbia Hall, Room 2370, 501 N Columbia Rd. Stop 7169, Grand Forks, ND 58202, USA.

acceptance, as their usage continues to grow, while still presenting several unique concerns. The findings will provide practitioners a guideline of what to address when communicating about voice assistants, or future AI devices, to potential users.

Literature review

Three factors influencing using voice assistants in TAM

Technological Factors

TAM is a popular theory used to predict acceptance and use of new technology (Davis, 1989) mainly derived from the Theory of Reasoned Action (TRA) (Ajzen & Fishbein, 1980; Fishbein & Ajzen, 1975). TAM posits that users accept or reject technology based on perceived usefulness (PU) and perceived ease of use (PEOU) (Davis, 1989). These variables have been used in recent applications and extensions of TAM on new technologies such as the telemedicine (e.g., Kamal, Shafiq, & Kakria, 2020), virtual reality (e.g., Manis & Choi, 2018) and smart appliances (Liu & Chou, 2020). Recent voice assistant studies have also used the traditional TAM variables (Song, 2019) or extensions including variables such as perceived enjoyment and innovativeness (Sorensen & Jorgensen, 2021). In the context of voice assistants, PU refers the degree to which the device makes life more efficient and helps carry out daily tasks, and PEOU refers to the degree to which using the device would be easy for the user. As voice assistants are new communication technology, it is appropriate to investigate why people accept and use them by adopting and extending TAM.

In this study, PU and PEOU, will be categorized as technological factors. Previous studies have found that PU and PEOU are drivers of attitudes toward using technology which ultimately impacts users' conscious behavior to use or reject the technology (Kamal et al., 2020; Manis & Choi, 2018). Generally, TAM studies found a positive relationship between attitudes toward using technology and intention to use it. Thus, this study suggests the following hypotheses:

H1: Attitudes toward use of voice assistants will be positively related to intention to use.

H2: PU will be positively related to intention to use voice assistants.

H3: PEOU will be positively related to PU.

H4: PU will be positively related to attitudes toward using voice assistants.

H5: PEOU will be positively related to attitudes toward using voice assistants.

Social Influence Factors

Social influence factors refer to norms that explain human behavior (Cialdini & Trost, 1998). Norms are defined as conceptions of what others think a person should do and what the person does (Smith, Terry, & Manstead, 2008). The social influence factors in this study include personal norms, what individuals believe their significant others (e.g., friends, family) think they should do and what they actually do, and societal norms, norms shared among general people (i.e., general public) in society (Smith et al., 2008). Based on previous explanations (Park & Smith, 2007), for voice assistants, personal norms can be formed when friends or family use and encourage using them. Whereas, social norms can be generated by information or data from news media or industry indicating many people buy voice assistants, and they recommend using them.

Personal and societal norms have been applied to predict people's attitudinal and behavioral responses (Park & Smith, 2007; Smith et al., 2008). Social norms approach indicated that personal and societal norms are significant antecedents of attitudinal and behavioral responses of a given behavior (Perkins, 2003). Additionally, Fishbein and Ajzen (1975) argued that to understand individuals' behaviors, it is important to consider norms as a social influence. In previous TAM studies, social influence has been considered as an additional factor to extend the original model (Maruping, Bala, Venkatesh, & Brown, 2017). It is predicted that perceived popularity and endorsement of using voice assistants by close others and society will generate a positive relationship with attitudinal responses toward use. Thus, the following hypotheses are suggested:

H6: Personal norms will be positively related to attitudes toward using voice assistants.

H7: Societal norms will be positively related to attitudes toward using voice assistants.

Risk Factors

Risk factors refer to potential threats users face when using technology. The major risk factors associated with voice assistants are privacy concerns and perceived security of the device. Research has demonstrated that these issues lead to negative outcomes such as lower brand loyalty (Hasan, Shams, & Rahman, 2021) and resistance to adopting and using the device (McLean & Osei-Frimpong, 2019; NPR, Edison Research Center, 2020).

Privacy concerns are defined as perceived vulnerability and perceived ability to control one's own personal information on the technology being used (e.g., gathering and usage of personal data) (Dinev & Hart, 2004), and perceived security is defined as perceptions regarding reliability of mechanisms of data transmission and storage (e.g., technical protection system) (Flavián & Guinalíu,

2006). By the definitions, privacy concerns are linked to the legal standards companies follow regarding personal data whereas perceived security is more related to technical aspects that guarantee privacy (Flavián & Guinalíu, 2006).

Regardless of the separate definitions between the privacy concerns and perceived security, studies have combined them into one variable (e.g., McCole, Ramsey, & Williams, 2010; Miyazaki & Fernandez, 2001). For example, although Flavián and Guinalíu (2006) defined privacy concerns and security as separate concepts, they ultimately combined them into a single construct. However, as technology has developed, users' awareness of privacy and security issues has grown. Now, technology users can likely recognize the difference between privacy and security. Instead of using a unidimensional variable, examining the two risk factors as separate constructs based on the conceptual differences would provide scholars and practitioners with more meaningful implications for the current state of technology concerns.

For voice assistants, privacy concerns refer to users feeling like their personal information is vulnerable (e.g., being listened to, watched, hacked) and having control over that personal information. Perceived security refers to technical aspects voice assistant companies have in place to protect users' personal information (e.g., passwords, voice recognition, secure networks). When using voice assistants, people consider both privacy concerns and perceived security (Orr & Sanchez, 2018).

Previous studies found a negative relationship between privacy concerns and a positive relationship between perceived security toward attitudinal responses of using new technology (McCole et al., 2010). It is expected that as people have more concerns about revealing their personal information, they have negative attitudes toward using voice assistants. On the other hand, if they have confidence about the protection of their information, they have more positive attitudes. Because of these reasons, this study has developed the following two separate hypotheses:

H8: Privacy concerns will be negatively related to attitudes toward using voice assistants.

H9: Perceived security will be positively related to attitudes toward using voice assistants.

Method

558 Amazon Mechanical Turk (MTurk) users were recruited for an online survey. Participants' ages ranged from 18 to 74 (M = 31.53, SD = 8.92). Male participants (59.3%, n = 331) outnumbered female participants (40.0%, n = 223). 55.4% (n = 309) of the participants were White or Caucasian, 25.8% (n = 144) were Asian,

and 7.3% (n = 41) were Black or African American. Overall participants were well-educated holding Bachelor's degree (57.7%, n = 322) or graduate/professional degree (13.4%, n = 75). Regarding the previous use of voice assistants, 94.3% (n = 526) have used voice assistants and 85.1% (n = 475) have owned one.

After consenting to participation, respondents were provided an explanation of voice assistants including their common uses (i.e., play music, shop online), the most popular examples (i.e., Apple's Siri, Amazon's Alexa, etc.), and the following definition, "Voice assistants are artificial intelligence technology that listen for specific words that activate them and then commands to complete from their users" (Gunkel, 2016; Hoy, 2018). Participants were directed to the main questionnaire which took about 15 minutes to complete.

Measures[1]

All measures used 7-point Likert scales ranging from *Strongly disagree* (1) to *Strongly agree* (7) except for attitudes. *Intention to use voice assistant* was measured with three items (e.g., I will use voice assistants within the foreseeable future.) modified from Manis and Choi (2018) (α = .88). *Attitudes toward using voice assistants* were measured by using five 7-point semantic differential scales (e.g., unfavorable – favorable) (Davis, 1989; Fishbein & Ajzen, 1975) (α = .93). *Perceived usefulness (PU)* was assessed by six items (e.g., Overall, I find using voice assistants advantageous.) borrowed from Davis (1989) (α = .91). *Perceived ease of use (PEOU)* was measured by using six items (e.g., I would find voice assistants easy to use.) borrowed and modified from Davis (1989) and Manis and Choi (2018) (α = .89). *Personal norms* were measured with six items (e.g., Most people whose opinion I value have used a voice assistant.) adopted from Park and Smith (2007) (α = .90). *Societal norms* were measured by using four items (e.g., A majority of people in the United States have used voice assistants.) from Park and Smith (2007) (α = .84). *Privacy concerns* were assessed by three items (e.g., I am concerned that voice assistants can obtain my personal information.) modified from Dinev and Hart (2004) (α = .85). *Perceived security* was measured with eight items (e.g., I think voice assistants show great concern for the security of any transactions.) borrowed from Flavián and Guinalíu (2006) (α = .94).

Results

Measurement model

A confirmatory factor analysis (CFA) was conducted by using AMOS 25 to validate the overall measurement constructs on adopting voice assistants of the hypothesized model in the context of the extend TAM model. The results of CFA confirmed two components of technological factors, two components of

social influence factors, and two components of risk factors, as well as attitudes and behavioral intentions. All standardized factor loadings in the measurement model were significant.[2]

Average variance extracted (AVE) is greater than .50 and composite reliability (CR) is greater than AVE for all constructs, indicating acceptable convergent validity. AVE is greater than maximum shared variance (MSV) for all constructs and the square root of AVE for each construct confirmed that it was higher than the correlations involving the construct, suggesting acceptable discriminant validity (Fornell & Larcker, 1981; Hair, Anderson, & Tatham, 2006) (see Table 1). As a result, all eight components of the hypothesized model were confirmed to be used as the measurement model in the hypothesized model.

Hypotheses testing[3]

To test the hypothesized model, AMOS 25 was used to run a structural equation modeling (SEM) with maximum likelihood estimation. As a result, overall fit indices for the hypothesized model indicated a good fit (GFI = .90; NFI = .93; CFI = .96; TLI = .96; IFI = .96; RMSEA = .04). Figure 1 shows the results of all hypothesized paths.

For technological factors, the results indicate that attitudes toward using voice assistants (β = .51, p < .001) and PU (β = .56, p < .001) were positively related to behavioral intentions, and that PEOU was positively related to PU (β = .75, p < .001). PU was positively related to attitudes toward using voice assistants (β = .53, p < .001), however PEOU was not (β = .01, n.s.). Thus, H1, H2, H3, and H4 were supported, whereas H5 was not. For social influence factors, personal norms were positively related to attitudes toward using voice assistants (β = .17, p < .01), whereas societal norms were not (β = .04, n.s.).

Table 1. Composite reliability (CR), average variance extracted (AVE), maximum shared variance (MSV), and correlation between constructs.

				Correlation between Constructs							
Constructs	CR	AVE	MSV	INT	ATT	PU	PEOU	PN	SN	PC	PS
INT	.88	.71	.65	.84*							
ATT	.93	.72	.65	.81	.85*						
PU	.92	.64	.61	.75	.78	.80*					
PEOU	.89	.56	.51	.71	.58	.70	.75*				
PN	.90	.60	.54	.63	.65	.69	.54	.78*			
SN	.83	.55	.54	.53	.54	.62	.58	.74	.74*		
PC	.85	.65	.19	.01	.03	.13	.18	.27	.44	.81*	
PS	.93	.64	.48	.56	.69	.69	.54	.66	.56	.05	.80*
Mean				5.34	5.39	5.04	5.35	4.82	4.98	4.74	4.73
SD				1.33	1.28	1.23	1.08	1.26	1.12	1.48	1.35

INT – Intention to Use Voice Assistant; ATT – Attitudes toward Using Voice Assistants; 1.35PU – Perceived Usefulness; PEOU – Perceived Ease of Use; PN – Personal Norms; SN – Societal Norms; PC – Privacy Concerns; PS – Perceived Security

*The numbers in the diagonal row are square roots of the average variance extracted.

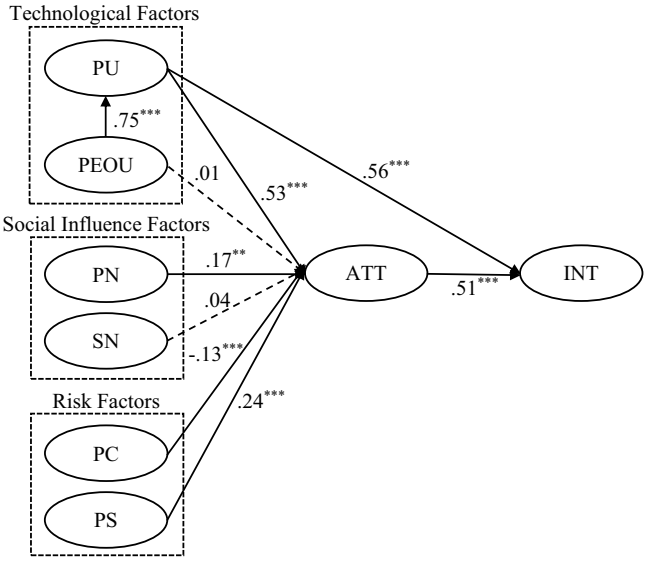

Figure1. The results of hypothesized model. $^{**}p < .01$, $^{***}p < .001$; significant path; non-significant pathχ^2 (661) = 1278.36, $p < .001$, χ^2/df = 1.93; GFI = .90; NFI = .93; CFI = .96;TLI = .96; IFI = .96; RMSEA = .04INT – Intention to Use Voice Assistant; ATT – Attitudes toward Using Voice Assistants; PU – Perceived Usefulness; PEOU – Perceived Ease of Use; PN – Personal Norms; SN – Societal Norms; PC – Privacy Concerns; PS – Perceived Security

Hence, H6 was supported while H7 was not. For risk factors, privacy concerns were negatively related to attitudes ($\beta = -.13$, $p < .001$) and perceived security was positively related to attitudes ($\beta = .24$, $p < .001$). Therefore, H8 and H9 were supported.

Discussion

This study extends TAM to focus on three factors that users consider when deciding to use voice assistants. Within technological factors, the results showed that when participants think voice assistants help their job performance, they generate more positive attitudes about using them, which in turn encourages their behavior intention. Moreover, when users think they can understand voice assistants quickly, they are more likely to consider them as useful technology. On the other hand, the effect PEOU had on attitudes differed from previous TAM studies (e.g., Davis, 1989; Manis & Choi, 2018). PEOU did not have a positive relationship with attitudes like hypothesized likely due to the increasing familiarity with technology. Although voice assistants are a new technology (NPR, Edison Research Center, 2020), adoption of new devices is easier than it once was due to increased exposure and the amount of readily available resources such as Google and YouTube. Therefore, PEOU is not a factor that affects people's

attitudes toward using voice assistants. The results of the social influence factors showed that societal norms are less important to users than personal norms implying that users' close network has a greater effect than general opinions. This result could be explained because voice assistants are personal devices. Moorthy and Vu (2014) found that users of voice activated mobile phone technology are more likely to use their device in private, not in public, and they are more likely to disclose non-private information to the device. Voice assistants are often used in private and can even be set up to control devices in your home or communicate to other family members. Users may prefer their close network's opinion about this technology because it is a sensitive and personal technology.

As predicted, privacy concerns had a negative influence on attitudes whereas perceived security had a positive relationship. It is noteworthy that, perceived security had a stronger influence toward attitudes. Because the relative strengths are different and show separate results, it is evident that participants understand privacy and security as different concepts.

Theoretically, this study suggests a TAM extension using three factors to explain technology acceptance. PEOU has always been a main variable of TAM (Davis, 1989). However, as users have become more familiar with technology, PEOU may not be as relevant as it once was. Therefore, TAM should be reevaluated to consider other variables for prediction of acceptance. The findings provide practical implications for the companies that produce voice assistants and for professionals who promote them. Because users have privacy concerns, voice assistant companies should include extra privacy and security measures. Communication professionals should express that voice assistants are useful, family-oriented, and secure putting emphasis on the product's usefulness, rather than how easy it is to use.

This study poses some limitations. First, there may be factors outside of the three in this study that affect voice assistant use. In order to further understand, a qualitative method such as focus group interviews could be conducted to gather other factors to test in the model. Second, by using an online survey panel like MTurk, participants might be more familiar with technology, including voice assistants, which may have affected the PEOU responses. A person who is already familiar with or who has already adopted a type of technology may not hesitate to use other new technologies. Future research could use a more traditional way of surveying to seek a variety of technology experience.

Notes

1. The Appendix A, which includes all measurement items used in this study, is available at https://mfr.osf.io/render?url=https%3A%2F%2Fosf.io%2Fpfhcq%2Fdownload.

2. The CFA results are available at https://mfr.osf.io/render?url=https%3A%2F%2Fosf.io%2Fn63xy%2Fdownload.
3. The Appendix B, which includes the summary of hypotheses testing, is available at https://mfr.osf.io/render?url=https%3A%2F%2Fosf.io%2Fmpj7h%2Fdownload.

Disclosure statement

There are no relevant financial or non-financial competing interests to report.

References

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behaviour*. Upper Saddle River, NJ: Prentice-Hall.

Cialdini, R. B., & Trost, M. R. (1998). Social influence: Social norms, conformity and compliance. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (pp. 151–192). New York, NY, US: McGraw-Hill.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. doi:10.2307/249008

Dinev, T., & Hart, P. (2004). Internet privacy concerns and their antecedents-measurement validity and a regression model. *Behaviour & Information Technology*, 23(6), 413–422. doi:10.1080/01449290410001715723

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.

Flavián, C., & Guinalíu, M. (2006). Consumer trust, perceived security and privacy policy. *Industrial Management and Data Systems*, 106(5), 601–620. doi:10.1108/02635570610666403

Fornell, C., & Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics. *Journal of Marketing*, 18(3), 382–388. doi:10.1177/002224378101800313

Gunkel, D. J. (2016). Computational interpersonal communication: Communication studies and spoken dialogue systems. *Communication+ 1*, 5(1), 1–20.

Guzman, A. L. (2019). Voices in and of the machine: Source orientation toward mobile virtual assistants. *Computers in Human Behavior*, 90(1), 343–350. doi:10.1016/j.chb.2018.08.009

Hair, J. F., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis* (6th ed.). Upper Saddle River, NJ: Prentice Hall.

Hasan, R., Shams, R., & Rahman, M. (2021). Consumer trust and perceived risk for voice-controlled artificial intelligence: The case of Siri. *Journal of Business Research*, 131, 591–597. doi:10.1016/j.jbusres.2020.12.012

Hoy, M. B. (2018). Alexa, Siri, Cortana, and more: An introduction to voice assistants. *Medical Reference Services Quarterly*, 37(1), 81–88. doi:10.1080/02763869.2018.1404391

Kamal, S. A., Shafiq, M., & Kakria, P. (2020). Investigating acceptance of telemedicine services through an extended technology acceptance model (TAM). *Technology in Society*, 60, 101212. doi:10.1016/j.techsoc.2019.101212

Liu, A. C., & Chou, T. Y. (2020). An integrated technology acceptance model to approach the behavioral intention of smart home appliance. *International Journal of Organizational Innovation*, 13(2), 95–118.

Manis, K. T., & Choi, D. (2018). The virtual reality hardware acceptance model (VR-HAM): Extending and individuating the technology acceptance model (TAM) for virtual reality hardware. *Journal of Business Research*. doi:10.1016/j.jbusres.2018.10.021

Maruping, L. M., Bala, H., Venkatesh, V., & Brown, S. A. (2017). Going beyond intention: Integrating behavioral expectation into the unified theory of acceptance and use of technology. *Journal of the Association for Information Science and Technology*, 68(3), 623–637. doi:10.1002/asi.23699

McCole, P., Ramsey, E., & Williams, J. (2010). Trust considerations on attitudes towards online purchasing: The moderating effect of privacy and security concerns. *Journal of Business Research*, 63(9–10), 1018–1024. doi:10.1016/j.jbusres.2009.02.025

McLean, G., & Osei-Frimpong, K. (2019). Hey Alexa . . . examine the variables influencing the use of artificial intelligent in-home voice assistants. *Computers in Human Behavior*, 99, 28–37. doi:10.1016/j.chb.2019.05.009

Miyazaki, A. D., & Fernandez, A. (2001). Consumer perceptions of privacy and security risks for online shopping. *Journal of Consumer Affairs*, 35(1), 27–44. doi:10.1111/j.1745-6606.2001.tb00101.x

Moorthy, A. E., & Vu, K. L. (2014). Privacy concerns for use of voice activated personal assistant in the public space. *International Journal of Human-Computer Interaction*, 31(4), 307–335. doi:10.1080/10447318.2014.986642

NPR, Edison Research Center. (2020). *The smart audio report spring 2020*. National Public Media. Retrieved July 21, 2021, from https://www.nationalpublicmedia.com/uploads/2020/04/The-Smart-Audio-Report_Spring-2020.pdf

Orr, D. A., & Sanchez, L. (2018). Alexa, did you get that? Determining the evidentiary value of data stored by the Amazon echo. *Digital Investigation*, 24, 72–78. doi:10.1016/j.diin.2017.12.002

Park, H. S., & Smith, S. W. (2007). Distinctiveness and influence of subjective norms, personal descriptive and injunctive norms, and societal descriptive and injunctive norms on behavioral intent: A case of two behaviors critical to organ donation. *Human Communication Research*, 33(2), 194–218. doi:10.1111/j.1468-2958.2007.00296.x

Perkins, H. W. (2003). The emergence and evolution of the social norms approach to substance abuse prevention. In *The social norms approach to preventing school and college age substance abuse* (pp. 3–17). San Francisco, CA: John Wiley & Sons.

Pitardi, V., & Marriott, H. R. (2021). Alexa, she's not human but . . . Unveiling the drivers of consumers' trust in voice-based artificial intelligence. *Psychology & Marketing*, 38(4), 626–642. doi:10.1002/mar.21457

Smith, J. R., Terry, D. J., & Manstead, A. S. R. (2008). The attitude-behavior relationship in consumer conduct: The role of norms, past behavior, and self-identity. *The Journal of Social Psychology*, 148(3), 311–333. doi:10.3200/SOCP.148.3.311-334

Song, Y. W. (2019). *User acceptance of an artificial intelligence (AI) virtual assistant: an extension of the technology acceptance model* (Doctoral dissertation).

Sorensen, K., & Jorgensen, J. J. (2021). "Hey Alexa, let's shop": Millennials' acceptance of voice-activated shopping. *International Journal of E-Services and Mobile Applications (IJESMA)*, 13 (1), 1–14. doi:10.4018/IJESMA.2021010101

Vimalkumar, M., Sharma, S. K., Singh, J. B., & Dwivedi, Y. K. (2021). 'Okay google, what about my privacy?': User's privacy perceptions and acceptance of voice based digital assistants. *Computers in Human Behavior*, 120, 106763. doi:10.1016/j.chb.2021.106763