The Uses and Gratifications Model of Voice Shopping

Jennifer Huh, Hye-Young Kim

University of Minnesota

Introduction

Voice shopping has become a buzzword as its popularity grows. According to a recent report, approximately 87.8 million of all adults in the U.S. have adopted voice assistants and the household adoption rate is to reach about 55% by 2022 (Voicebot, 2019). The growing popularity of voice assistants draws upon their functional characteristics. Assisted by Artificial Intelligence (AI), voice assistants can function as decision aid tools and personal assistants that help consumers in their daily decision-making (Mari et al., 2020).

Past studies based on the Technology Acceptance Model (TAM) or related theories (e.g., McLean & Osei-Frimpong, 2019; Yang & Lee, 2019) provided preliminary accounts for consumers' adoption intentions. Also, another research stream based on social response theory or human-machine interaction (HMI) focused on the relational aspects (e.g., Moriuchi, 2021) and provided novel insights. However, scant research provided a comprehensive approach to understand the underlying psychological mechanisms of AI-enabled voice assistant usage and voice shopping behavior.
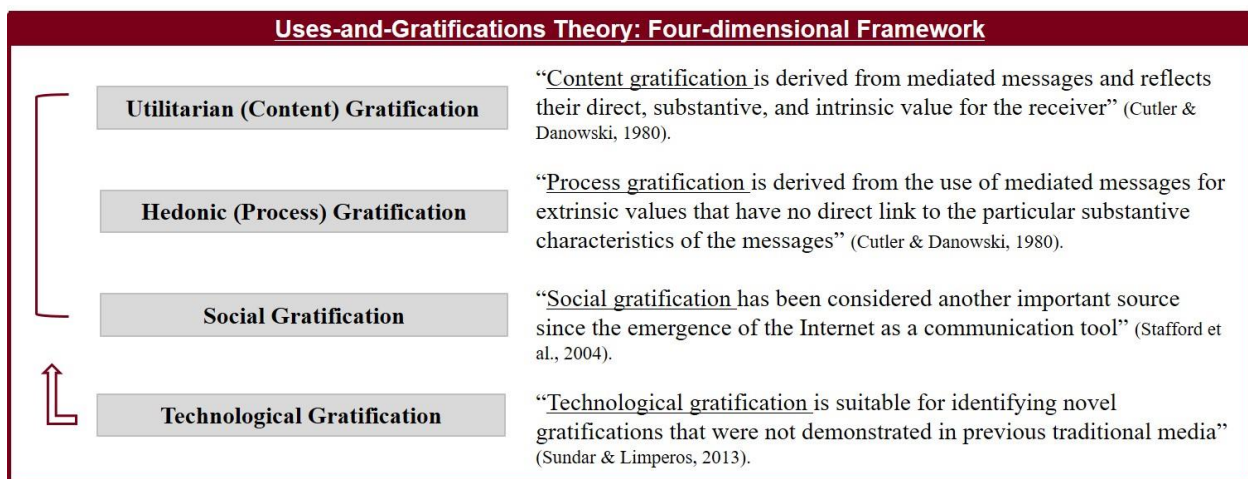
Therefore, this paper aims to investigate the voice shopping phenomenon. Specifically, the purpose of this research is to examine; (a) whether different gratification dimensions a voice assistant' users experience have differential influence on overall satisfaction and (b) whether overall satisfaction leads to fashion product purchases through a voice assistant.
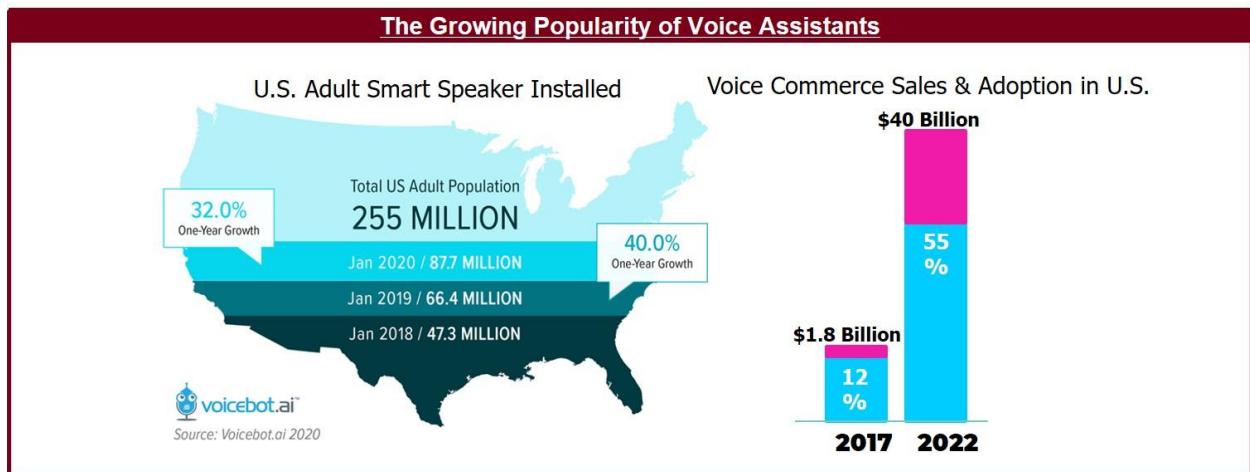
Literature Review and Hypotheses

This research adopted the uses and gratifications theory (UGT) as a theoretical basis. First, the UGT is based on the user-centric approach, which is useful for identifying consumers' context-specific motivations for using voice assistants (Katz et al., 1973). Second, its explanatory power is relevant for investigating a new media (Liu, 2015). Finally, its psychological perspective helps researchers explain the underlying factors for voice shopping (Palmgreen et al., 1985).

Specifically, this research adopted the four-dimensional framework from the UGT. Previous gratifications literature mainly identified gratifications from three types of sources, which are utilitarian, hedonic, and social-oriented aspects (Cutler & Danowski, 1980; Stafford et al., 2004). However, this research incorporates the fourth source, technological gratification, as it helps researchers to identify a new gratification dimension from using new media, such as AI-enabled autonomous devices (Sundar & Limperos, 2013).

Uses-and-Gratifications Theory: Four-dimensional Framework

Utilitarian (Content) Gratification	"Content gratification is derived from mediated messages and reflects their direct, substantive, and intrinsic value for the receiver" (Cutler & Danowski, 1980).
Hedonic (Process) Gratification	"Process gratification is derived from the use of mediated messages for extrinsic values that have no direct link to the particular substantive characteristics of the messages" (Cutler & Danowski, 1980).
Social Gratification	"Social gratification has been considered another important source since the emergence of the Internet as a communication tool" (Stafford et al., 2004).
Technological Gratification	"Technological gratification is suitable for identifying novel gratifications that were not demonstrated in previous traditional media" (Sundar & Limperos, 2013).

To provide a conceptual background for each gratification dimension, content gratification is derived from mediated messages and reflects their direct, substantive, and intrinsic value for the receiver (Cutler & Danowski, 1980). This study adopted life efficiency, "the degree to which the voice assistant helps the user to complete daily tasks such as searching for information, providing event reminders, and placing an order efficiently based on voice commands (Rauschnabel, 2018)," as the source for content/utilitarian gratification.

Process gratification is derived from the use of mediated messages for extrinsic values that have no direct link to the particular substantive characteristics of the messages (Cutler & Danowski, 1980). This study adopted entertainment, "the degree to which the voice assistant offers a pleasurable experience through its usage (Huang, 2008; Zhang et al., 2011)," as the source for process/hedonic gratification.

Social gratification has been considered another important source since the emergence of the Internet as a communication tool" (Stafford et al., 2004). This study adopted social presence,

"the degree to which the voice assistant allows the user to feel a sense of interpersonal interaction through its usage" (McLean & Osei-Frimpong, 2019; Gefen & Straub, 2003)," as the source for social gratification.

Finally, technological gratification is suitable for identifying novel gratifications that were not demonstrated in previous traditional media (Sundar & Limperos, 2013). This study adopted affordance, "the degree to which the voice assistant enables the user to readily recognize the actions it can perform easily (Bae et al., 2016; Bae, 2018)," as the source for technological gratification.

- **H1:** A high level of life efficiency will lead to a high level of overall satisfaction.

- **H2:** A high level of entertainment will lead to a high level of overall satisfaction.

- **H3:** A high level of social presence will lead to a high level of overall satisfaction.

- **H4:** A high level of affordance will lead to a high level of overall satisfaction.

- **H5:** A high level of overall satisfaction will lead to a high level of fashion product purchases through a voice assistant.

- **H6:** Overall satisfaction will mediate the relationship between a) life efficiency, b) entertainment, c) social presence, d) affordance, and fashion product purchases through a voice assistant.


Methods

This research was based on a self-administered online survey method. The research setting was tested in the context of Amazon voice shopping by recruiting the actual users of Alexa. The participants were recruited using Pollfish. Pollfish is a company that recruits respondents in a real-time through their mobile-application developers and utilizes machine learning techniques to eliminate poor quality respondents (https://pollfish.com). Amazon M Turk service is often raised concerns that workers tend to be more male, White, educated, Democrat, liberal, and younger than the overall adult US population, thus not representing an actual random sampling (Sheehan, 2018). However, Pollfish is known to best align with a non-probability based survey (Goel, Obeng & Rothschild, 2015). Thus, this research recruited respondents from Pollfish, where they were directed to the survey through the mobile application advertising page, then rewarded upon the completion of the survey.

A total of 166 responses was collected. The majority of the participants were male (55.4%), in their 30s and 40s (60.2%), and Caucasian (71.7%), with household income over $35K and below $110 K (53.1%). The measurement items were adopted from previous literature.

Result

Partial Least Squares (PLS) modeling was used due to the small sample size and the exploratory purpose of the present research. The measurement model was tested by assessing Cronbach's alpha, average variance extracted (AVE), composite reliability (CR), discriminant validity, and the results were all satisfactory. The evaluation of structural model was tested by estimating the effect sizes (f^2), predictive relevance (q^2), R^2 (satisfaction = .533, purchase = .315), goodness of fit (srmr = .079). The hypothesis testing revealed that H1 (β= .430, t= 4.848, p <.001), H2 (β= .262, t= 2.585, p <.01), and H5 (β= .561, t= 8.91, p <.001) were supported. Mediation analysis also revealed that H6$_a$ (β= .248, t= 3.872, p <.001) and H6$_b$ (β= .152, t= 2.457, p <.01) were supported.

Figure 1. Results of the structural model
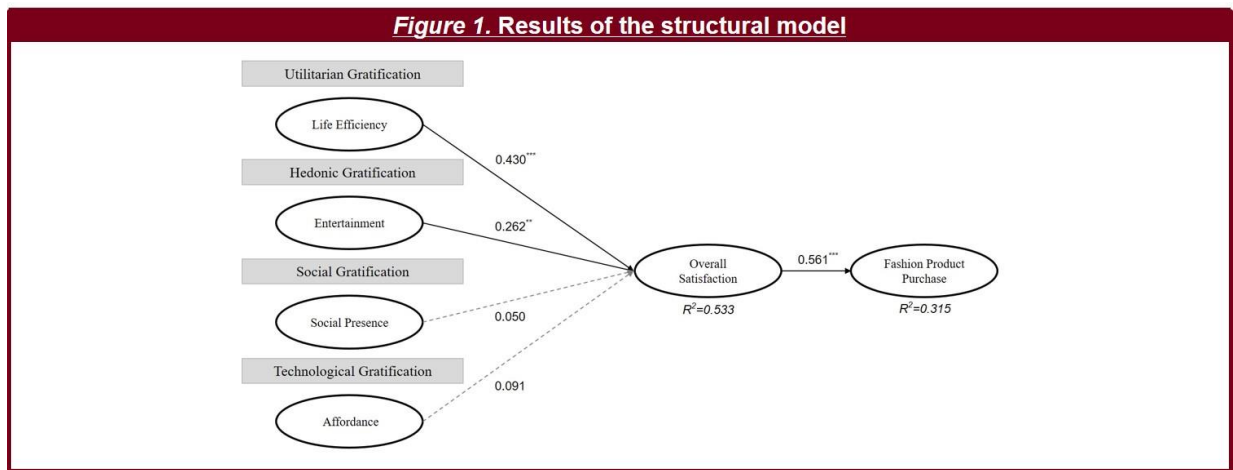
Note: ***p < .001, **p < .01.

PLS-SEM Results

								*Lower-Limit/Upper-Limit	
Hypothesis	Relationship	Std Beta	Std Error	[t-value]^	Decision	f2	q2	95% CI LL	95% CI UL
H4	affordance -> satisfaction	0.091	0.106	0.846	-	0.006	-0.003	-0.087	0.271
H1	efficiency -> satisfaction	0.43	0.089	4.848***	Supported	0.165	0.078	0.27	0.56
H2	entertainment -> satisfaction	0.262	0.101	2.585**	Supported	0.07	0.035	0.101	0.436
H3	social -> satisfaction	0.05	0.065	0.798	-	0.004	-0.003	-0.063	0.152
H5	satisfaction -> purchase	0.561	0.063	8.91***	Supported			0.458	0.659

Conclusion

This study investigated the voice shopping phenomenon based on the uses and gratifications perspective. The study results revealed that H1, H2, H5, and H6a, H6b were supported.

However, H3, H4, and other related mediating process were rejected. Although social presence (H3) and affordance (H4) did not show a significant result, the research finding indicates the importance of incorporating utilitarian and hedonic gratification in using AI-enabled voice assistants.

Understanding important gratification dimensions sought from voice assistants can guide fashion retailers in fine-tuning their voice commerce services. For example, the significant effect of life efficiency highlights the managerial importance of helping shoppers perform tasks at a much faster speed and with less cognitive effort (Rauschnabel, 2018). Fashion retailers could benefit from enhancing efficiency on their voice commerce services; just as H&M launched a shopping guide and Estee Lauder introduced custom-made skincare solutions through voice activation (Ball, 2020). Also, the findings of this study suggest that voice shopping can certainly fit the hedonic consumption perspective (Holbrook & Hirschman, 1982). This is consistent with the current trend that voice assistant developers are accelerating the creation of content that feeds to voice assistant ecosystems ranging from voice games and storytelling to education and entertainment (Modev, 2020). Prior research suggests that the voice modality of assistants can increase the positive attitude toward the device as users perceived greater human likeliness (Cho, Molina, & Wang, 2019). However, this study could not confirm this social presence effect. With respect to affordance, the significant finding of the study implies that discovering voice applications is still a challenge and thus finding out what shopping options exist on voice and understanding how to access them can be a challenge for users (Simms, 2019). This finding contradicts with the contention that the voice and audio-based modality becomes a powerful affordance influencing the interaction between users and digital media, which in turn affects the users' perceptions associated with the content and the media platform (Cho, 2019).

This study has limitations. First, the sample size was small. This study recruited the respondents from Pollfish, a relatively new platform that identifies participants from company's partnered mobile application users. Although some studies appreciated its innovative approach using machine-learning system and mobile users' fast responses rate (e.g., Papagiannaki et al., 2021), a caution is needed for using this service as some participants could have decided to participate in a survey for acquiring intensive (e.g., game money).

Finally, the following future lines of research can conduct a study that elaborates the psychological process of voice shopping behavior. For example, a study examining how each unique functional characteristics of AI-enabled voice assistants contribute to consumers' relationship building could add novel insights and provide practical implications for retailers on how to show presence in voice channel.

References

Bae, M. (2018). Understanding the effect of the discrepancy between sought and obtained gratification on social networking site users' satisfaction and continuance intention. Computers in Human Behavior, 79, 137-153. https://doi.org/10.1016/j.chb.2017.10.026

Bae, Y. H., Jun, J. W., & Hough, M. (2016). Uses and gratifications of digital signage and relationships with user interface. Journal of International Consumer Marketing, 28(5), 323-331. https://doi.org/10.1080/08961530.2016.1189372

Ball, E. (2020). Is voice search within fashion E-commerce the future? *Prestige Recruitment Group*. Retrieved from https://prestigerecruitmentgroup.com

Cho, E. (2019, May). Hey Google, Can I Ask You Something in Private?. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1-9).

Cho, E., Molina, M. D., & Wang, J. (2019). The effects of modality, device, and task differences on perceived human likeness of voice-activated virtual assistants. *Cyberpsychology*, *Behavior, and Social Networking*, 22(8), 515-520.

Cutler, N. E., & Danowski, J. A. (1980). Process gratification in aging cohorts. *Journalism Quarterly*, 57(2).

Gefen, D., & Straub, D. (2003). Managing user trust in B2C e-services. e-Service, 2(2), 7-24.

Goel, S., Obeng, A., & Rothschild, D. (2015). Non-representative surveys: Fast, cheap, and mostly accurate. manuscrit non publié.

Hirschman, E. C., & Holbrook, M. B. (1982). Hedonic consumption: emerging concepts, methods and propositions. *Journal of marketing*, 46(3), 92-101.

Huang, E. (2008). Use and gratification in e-consumers. Internet Research, 18(4), 405-426. https://doi.org/10.1108/10662240810897817

Katz, E., & Gurevitch, M. Hass (1973). On the use of the mass media for important things. *American Sociology Review*, 50, 340-359.

Liu, W. (2015). A historical overview of uses and gratifications theory. *Cross-Cultural Communication*, 11(9).

Mari, A., Mandelli, A., & Algesheimer, R. (2020, July). The evolution of marketing in the context of voice commerce: A managerial perspective. In *International Conference on Human-Computer Interaction*. Springer, Cham. 405-425.

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. https://doi.org/10.1016/j.chb.2019.05.009

Modev. (2020, June 18). Learn "How Voice Technology is Shaping Education and Entertainment" *CISION*. Retrieved from https://www.prnewswire.com

Moriuchi, E. (2021). An empirical study on anthropomorphism and engagement with disembodied AIs and consumers' re-use behavior. Psychology & Marketing, 38(1), 21-42. https://doi.org/10.1002/mar.21407

Palmgreen, P., Wenner, L. A., & Rosengren, K. E. (1985). Uses and gratifications research: The past ten years. In *Media gratifications research: Current perspectives*, 1, 1-37.

Papagiannaki, K., Diakakis, M., Kotroni, V., Lagouvardos, K., & Papagiannakis, G. (2021). The Role of Water Depth Perception in Shaping Car Drivers' Intention to Enter Floodwaters: Experimental Evidence. Sustainability, 13(8), 4451.

Rauschnabel, P. A. (2018). Virtually enhancing the real world with holograms: An exploration of expected gratifications of using augmented reality smart glasses. Psychology & Marketing, 35(8), 557-572. https://doi.org/10.1002/mar.21106

Rauschnabel, P. A. (2018). Virtually enhancing the real world with holograms: An exploration of expected gratifications of using augmented reality smart glasses. *Psychology & Marketing*, 35(8), 557-572.

Sheehan, K. B. (2018). Crowdsourcing research: data collection with Amazon's Mechanical Turk. Communication Monographs, 85(1), 140-156. https://doi.org/10.1080/03637751.2017.1342043

Simms, K. (2019). How voice assistants could change the way we shop. *Harvard Business Review*. Retrieved from https://hbr.org/

Stafford, T. F., Stafford, M. R., & Schkade, L. L. (2004). Determining uses and gratifications for the Internet. *Decision sciences*, 35(2), 259-288.

Sundar, S. S., & Limperos, A. M. (2013). Uses and grats 2.0: New gratifications for new media. *Journal of Broadcasting and Electronic Media*, 57(4), 504-525.

Voicebot (2019). Smart speaker consumer adoption report. *Voice bot*. Retrieved from https://voicebot.ai/

Yang, H., & Lee, H. (2019). Understanding user behavior of virtual personal assistant devices. Information Systems and e-Business Management, 17(1), 65-87. https://doi.org/10.1007/s10257-018-0375-1

Zhang, Y., Tang, L. S. T., & Leung, L. (2011). Gratifications, collective self-esteem, online emotional openness, and traitlike communication apprehension as predictors of Facebook uses. Cyberpsychology, Behavior, and Social Networking, 14(12), 733-739. https://doi.org/10.1089/cyber.2010.0042