

Artificial Intelligence-based Voice Assistants as Social Media Platform: Anticipated Effect on Brands During Shopping

Alex Mari

Keywords

Voice assistant, social media, voice commerce, brand management, platform ecosystem.

Abstract

Artificial intelligence (AI) technologies enable voice assistants (VA) such as Amazon Alexa and Alibaba Genie to perform activities that resemble cognitive functions associated with the human mind (Norvig, 2011). While performing complex tasks with consumers, VAs become more human-like exchange partners. As these AI-enabled devices learn consumer preferences and habits, they introduce biases and heuristics likely to affect individual and collective shopping behavior while posing new opportunities and threats for managers. Why should managers monitor the rapid diffusion of this novel 'social media' platform? What are the anticipated consequences for brands in the shopping context? This interpretative study uses in-depth interviews with elite informants (N = 31) to address these questions. Executives and experts predict that VAs assume a central relational role in the consumer market and progressively mediate market interactions. Simultaneously, they expect VA's intrinsic social nature to influence consumption decisions increasingly. These fast-changing market dynamics within the context of voice commerce may have severe consequences for consumer brands. The main identified threats for managers and their implications are discussed.

Introduction

The prolific environment for AI-based devices is expected to develop at an exponential rate while altering consumer decision-making and posing new challenges for managers (Davenport et al., 2020). AI enables, among others, VAs to systematically collect consumer information, identify patterns, and predict future behaviors (Mari et al., 2020). The physical placement of these devices at the core of consumers' domestic life allows for repeated, ongoing interactions that fulfill functional and social needs (Ammari et al., 2019; McLean et al., 2021). As such, VAs are expected to become a social media hub influencing individual and collective shopping decisions.

Due to rapid adoption and disruptive potential in buying dynamics, VAs are considered a transformative consumer media in academic (Dellaert et al., 2020), business (Dawar & Bendle, 2018), and industry (eMarketer, 2019) research. However, little is known about managers' perceptions of how AI-based VAs can alter consumer judgment and behavior toward brands during shopping practices (Mari & Algesheimer, 2021a) it is unclear 'when' voice assistants are capable of gaining trust and 'how' the development of such a trusted relationship affects decisions. This research explores the effect of trusting beliefs towards voice assistants on decision satisfaction through the indirect effect of consideration set size (n. of options). Although exemplary research on media management and innovation offer insights that are likely transferable to voice

assistants, the peculiarities of this technology require new theories that are not yet fully developed (Kumar et al., 2016). This study adopts a managerial perspective to shed light on the potential impact that the diffusion of shopping-related VAs has on consumer brands. The first section provides a review of the most relevant literature. Next, the conceptualization of the dual media role of VAs (agentic and mediator) is discussed before outlining the anticipated consequences for consumer brands.

Artificial Intelligence-based Voice Assistants as Social Media

Defined as 'voice-based interfaces that can actively guide consumer decisions on the basis of artificial intelligence' (Dellaert et al., 2020, p.2), VAs can naturally converse with users, contextually elaborate requests, and dynamically increase their knowledge while learning from mistakes (Mari & Algesheimer, 2021b). AI-based VAs leverage advanced AI techniques, such as automatic speech recognition and natural language understanding, to engage in dialogues with an individual through human language. These smart devices operate as a personal concierge or shopper while executing a large variety of user requests in real-time (Hoy, 2018; Lee & Choi, 2017). During interactions with consumers, VAs can process (and store) a considerable amount of input data and thus incrementally understand its consumers and context. In particular, VAs provide adaptive responses by processing contextual clues, such as the user's identity and the device's

location, while learning to correct unsatisfactory interactions shortly after they occur (Huang & Rust, 2018; Knotte et al., 2019; Sarikaya, 2017) constituting a major source of innovation, yet threatening human jobs. We develop a theory of AI job replacement to address this double-edged impact. The theory specifies four intelligences required for service tasks—mechanical, analytical, intuitive, and empathetic—and lays out the way firms should decide between humans and machines for accomplishing those tasks. AI is developing in a predictable order, with mechanical mostly preceding analytical, analytical mostly preceding intuitive, and intuitive mostly preceding empathetic intelligence. The theory asserts that AI job replacement occurs fundamentally at the task level, rather than the job level, and for “lower” (easier for AI. A clearer picture of the interlocutors - individuals and other smart objects - increases a VA’s ability to individualize, personalize, and contextualize voice-based requests. While proactively learning consumer preferences and habits, VAs improve predictive accuracy and increasingly exhibit independent behavior (Rahwan et al., 2019). As such, they are set to influence consumer decision-making in new ways (Schmitt, 2019) the Industrial Revolution and subsequent technology-driven economic transformations have radically changed how people live and work, and how consumers buy and use goods and services. We are now in the midst of another major technological transformation—the digital revolution. In *Being Digital*, Nicholas Negroponte, the visionary founder of the MIT Media Lab, used the metaphor of a shift “from atoms to bits” to characterize digitization (Negroponte 1995.

While performing complex tasks for and with consumers, VAs are also becoming more human-like exchange partners. The relationship between humans and VAs is fundamentally social because it activates emotional, cognitive, and behavioral reactions that are usually found in human-to-human relationships (Reeves & Nass, 1996). The field of human-machine interaction (HCI) largely accepts that people treat technological artifacts not only as a medium but also as a social actor (Chérif & Lemoine, 2019; Nass & Moon, 2000; Wang et al., 2007). Per the ‘computers are social actors’ paradigm, individuals mindlessly attribute human-like characteristics to agents, such as VAs, and apply human attributes (e.g., personality traits, stereotypes, and norms) when interacting with them (Reeves & Nass, 1996). When an artificial agent is considered capable of agentic communication, the consumer’s sense of human warmth and sociability (called social presence) increases (Gefen & Straub, 2004, Short et al., 1976). In an online shopping context, conversational agents can induce social presence with a strong positive effect on the user’s trusting beliefs, satisfaction and ultimately willingness to accept product recommendations (Holzwarth et al., 2006; Lu et al., 2016; Qiu & Benbasat, 2009).

Social response theory postulates that individuals tend to consider a social entity an agent if it possesses even minimal indices of similarity with humans and shows scarce automation capabilities (Nass & Steuer, 1993). However, social responses appear to be more assertive when anthropomorphized agents display human traits and characteristics, such as facial expressions, body gestures, or a voice (Araujo,

2018; Go & Sundar, 2019). In particular, the voice leads to a 'persona effect' (Lester et al., 1997, p.359) because it powerfully indicates the presence of another person (Nass & Gong, 2000). Similar to humans, VAs assume a persona ('I') to refer to themselves, and they react to interlocutors when their name is called. AI-powered VAs also retain relevant facts during the conversation to give a sense of continuity to subsequent interactions.

The use of speech in smart objects also generates physiological and affective arousal because the voice incorporates rich non-verbal cues through varying tones, intonation, speed, and emphasis on words (Scherer, 2003). People can make inferences regarding emotions, attitudes (Giles, 1970), social status (Nass & Brave, 2005), and personality (Aronovitch, 1976) from vocal characteristics. As such, oral interaction with machines is perceived more favorably by users than text-only interaction because it allows for more natural and effective dialogue (Nass & Scott, 2005) in the shopping context (Qiu & Benbasat, 2009). Reeves and Nass (1996) argue that the human brain did not evolve fast enough to elaborate on the ontological nature of an agent. In that sense, when a machine is involved in real-time language production that allows engagement in unstructured dialogues with reciprocal responding, users subconsciously adopt the same social rules and heuristics that are practiced in their interpersonal relationships (Nass & Moon, 2000). Overall, researchers agree that the use of software-based speech on smart devices, such as Amazon Alexa or Alibaba Genie, can significantly strengthen the interpersonal nature of HCIs (Rhee & Choi, 2020).

Humanizing Voice Assistants

Voice assistants are 'always on' devices that can process (or even automate) tasks with a simple voice command and without the user having to provide repetitive personal information such as address or credit card details. Different from other digital platforms, VAs do not favor multitasking activities as they are designed to process one request at a time and on a turn-by-turn basis. Also, as the interaction is limited to auditory cues, information is delivered sequentially with a higher mnemonic effort. Such characteristics unique to VAs require a new set of interaction rules modeled after the active (and proactive) nature of these AI-enabled devices (Schmitt, 2019) the Industrial Revolution and subsequent technology-driven economic transformations have radically changed how people live and work, and how consumers buy and use goods and services. We are now in the midst of another major technological transformation—the digital revolution. In *Being Digital*, Nicholas Negroponte, the visionary founder of the MIT Media Lab, used the metaphor of a shift "from atoms to bits" to characterize digitization (Negroponte 1995).

In the context of VAs and beyond, AI is rapidly evolving from pure mechanical capabilities to analytical capabilities based on intuition, with the next frontier represented by empathetic capabilities, essential to explore and understand human emotions (Huang & Rust, 2018) constituting a major source of innovation, yet threatening human jobs. We

develop a theory of AI job replacement to address this double-edged impact. The theory specifies four intelligences required for service tasks—mechanical, analytical, intuitive, and empathetic—and lays out the way firms should decide between humans and machines for accomplishing those tasks. AI is developing in a predictable order, with mechanical mostly preceding analytical, analytical mostly preceding intuitive, and intuitive mostly preceding empathetic intelligence. The theory asserts that AI job replacement occurs fundamentally at the task level, rather than the job level, and for “lower” (easier for AI. Thanks to AI advancements, human–VA interaction will increasingly resemble the human-to-human relationship. Voice assistant manufacturers have recently put in place a series of activities to maximize the realism of voice-based conversations. For instance, Google’s team has launched ‘Duplex’, an automated voice assistant capable of generating realistic speech almost indistinguishable from a human voice, posing critical ethical questions on the bot’s self-disclosure strategies (Ishowo-Oloko et al., 2019). At the same time, Amazon launched the Alexa Prize, a global competition for university students focused on creating a ‘social bot’ that converses coherently and engagingly with humans on popular topics such as sports, politics, and entertainment (Ram et al., 2018).

Built to mimic natural human-to-human interactions, Amazon’s text-to-speech technology enables owners of third-party applications to create speech that sounds more natural and intuitive. For instance, Alexa can respond in

a happy/excited tone when a customer answers a trivia question correctly or in a disappointed/empathetic tone when the answer is wrong. Internal Amazon studies reported an overall satisfaction increase with the voice experience by 30% when Alexa responded with emotions (Gao, 2019). Additionally, Alexa can be coded to respond in a speaking style that is more suited for a specific type of content, such as news and music. By changing characteristics of speech such as intonation, pauses and emphasis on keywords, Matthew, Joanna, and Lupe voices – the available personas – can either sound similar to a TV news anchor, or less formal, like a friend. In internal ‘blind listening’ tests, Amazon reported that the styles were perceived up to 84% more natural than Alexa’s standard voice (Gao, 2020). Furthermore, the latest Alexa speech capabilities allow developers to mix multiple voices with sound effects (e.g., fading in and out) and other audio clips to create a more natural-sounding experience (Gao, 2020). The ability of VAs to display emotions through voice and engage in casual jokes makes them pleasant conversational partners (Han & Yang, 2018). As captured in the developments above, the importance of realism maximization resides in greater user satisfaction (Von Der Pütten et al., 2010). Anthropomorphic cues and AI technologies may jointly contribute to strengthening the feeling of social presence with an effect on the perceived credibility and competence of voice assistants. With the distinction between humans and machines becoming increasingly blurred, AI-based VAs are expected to substan-

tially alter the decision-making process throughout the entire consumer journey, from product search to repurchase (e.g., Mandelli, 2018).

Interactional Characteristics of Shopping-related Voice Assistants

In contrast to traditional media, voice touchpoints emphasize a bidirectional interaction with consumers. VAs are designed to process one request at a time and on a turn-by-turn basis to decrease the speech recognition error rate coming from a possible voice overlap (Hansen, 1996) several causes for recognition performance degradation are explored. It is suggested that recent studies based on a Source Generator Framework can provide a viable foundation in which to establish robust speech recognition techniques. This research encompasses three inter-related issues: (i. This style of interaction represents a radical difference compared to sensorially richer devices like computers or smartphones, which present multiple pieces of information on a screen concurrently. As such, voice channels present both challenges and opportunities for the diffusion of commerce practices.

On the positive side, e-commerce has paved the way for voice shopping (Labecki et al., 2018). With the rise of the Internet, users have learned to deal with a combination of social, cultural, economic, and technical barriers. In doing so, they overcame the initial diffidence of buying without directly seeing, touching, or smelling an object. Voice technologies further limit

the users' senses; besides, consumers are asked to make shopping decisions without browsing photos, videos, or any other animated content. Another celebrated feature of voice shopping is the ease of making low-involvement purchases. With a simple 'yes' and without providing additional information such as credit cards or address details, VAs can conveniently process orders.

On the negative side, an effortless decision-making process does not guarantee an optimal level of consumer satisfaction. Shopping-related VAs offer a limited set of items for each product category based on their understanding of the consumer and context. This simplified representation of the marketplace reduces consumers' visibility of product alternatives and emphasizes the critical role of ranking algorithms. The algorithm that ranks the information represents a 'black box' for the VA user and the brand owner (Voosen, 2017, p.22). Furthermore, such visual limitations may increase brand polarization while enhancing the risk of the so-called 'filter bubble' or 'echo-chamber' effects (Colleoni et al., 2014, p.317).

The terms 'voice commerce' and 'voice shopping' refer to the transactional act of placing an order directly or through marketplaces and to the technical capabilities and communication activities that allow users to search for a product, listen to reviews, add items to a list, purchase goods or services, track the order, access customer service, etc. (Mari & Algesheimer, 2021b). The number of consumers who have completed at least one purchase through a smart speaker widely varies among product categories. A report suggests that 21% of U.S. smart speaker owners have purchased

entertainment such as music or movies, 8% household items, and 7% electronic devices (eMarketer, 2019). Moreover, the exponential growth of app marketplaces with over 100,000 voice 'skills' available on the Alexa Skills Store, and 10,000 'actions' on Google Home, provides infinite interaction opportunities between brands and consumers (Sterne, 2017).

In this context, managers came to realize that VAs already possess the technical capabilities to 'lead' the interaction with consumers, from activating passive users to automating product purchases.

Methodology

This interpretative study uses one-to-one in-depth interviews to capture the experiences, feelings, perceptions, and knowledge of AI-aware corporate executives and international consultants. To favor the process of data collection standardization and maximize reliability, validity and generalization, interviews were semi-structured. A total of 31 elite informants participated in the study in December 2018 (Solarino & Aguinis, 2020). During the interview process, theoretical perspectives were not employed to facilitate the emergence of insights (Avis, 2003). Interviews were audio-taped and transcriptions analyzed adopting an inductive line-by-line coding approach following grounded theorizing. This method was chosen as it employs inductive theory building with explicit use of coding procedures (e.g., Strauss & Corbin, 1990). Following a constant comparative data analysis (Glaser & Strauss, 1967), levels of abstraction were generated directly from the data

and refined through additional data until key themes reaching saturation. Using NVivo 12 for Mac, codes were grouped into themes and then re-evaluated to ensure that they reflect data extracts. Through conceptualization, relationships among categories and sub-categories were established.

Findings and discussion

The following paragraphs describe the emerging conceptual nodes. First, an overview of the dual media role of VAs – agentic and mediator – is provided. Second, the anticipated negative effects for consumer brands coming from the diffusion of VAs are discussed.

The Agentic Role of Voice Assistants

VAs are perceived by managers to have an agentic role as they attempt to predict which items a target user likes based on expressed preferences or implicit behaviors (Shen, 2014). The informant, Jim Sterne, Emeritus Director of the Digital Analytics Association (DAA), explains:

There is a brand of soap that my wife loves. One day Alexa says, 'Hey, you buy this all the time, why don't you subscribe?'. Now, we have a subscription to soap, and every six months we get a bunch. If we have more than we need, we adjust the delivery frequency. This product automatically shows up, and we are definitely going to buy the same brand. We are locked in.

According to the finding, this form of recommendation system may replace traditional decision-making when consumers feel time constraints or recognize the referrer as a particularly knowledgeable source (Olshavsky & Granbois, 1979). Throughout the collection of significant volumes of personal and behavioral information, VAs can push users to automate repurchase, for instance, via 'subscribe & save' promotional activities, which are increasingly popular on e-commerce websites. According to André et al. (2018), this power of attorney toward VAs comes at the expense of higher-order psychological processes such as emotions and moral judgments. In the context of purchase automation, consumers might have aspirational preferences that differ from the preferences suggested by their past behavior. These meta preferences, also called preferences over preferences (Jeffrey, 1974), are apparent in the case of an environmentally aware person who wants to use less bottled water but is regularly reminded to buy plastic bottles. The inherent tension between the actual-self and the ideal-self represents a boundary for those consumers who follow VAs' suggestions to automate repurchases.

The Mediator Role of Voice Assistants

Due to its central role in a complex business network, VAs do not consider users as the only stakeholders benefiting from the recommended outcome. The strategic goals of the retailer, merchant,

advertiser, and VA itself may differ from those of end-users. As such, the user is not the sole focus of a VA recommendation. For instance, a VA might recommend a private label over a consumer brand following the retailer's objective to swiftly grow its shares in a specific product category. Thus, the interests of several parties must coexist (Abdollahpouri et al., 2020).

Informants believe that VAs are increasingly mediating consumer interactions with the market with an effect on the path to purchase dynamics. While functioning as a salesperson, VAs are redefining relationships among consumers, brands, and retailers.

Also, due to the high media speculation of the negative impact that shopping-related VAs have on brands, managers feel threatened by the bargaining power shifting in favor of VA manufacturers (Dawar & Bendle, 2018). In the case of Amazon Alexa, the VA manufacturer is also the retailer behind the most advanced voice shopping functionality, accounting for nearly 45% of the total U.S. retail e-commerce (eMarketer, 2019).

Consequences of the Diffusion of Voice Commerce on Brands

Findings show six main areas of concern in informants connected to the rise of voice shopping. Tab.1 summarizes the propositions, offers a general statement and provides evidence with an exemplary quote from the conducted interviews.

Proposition	General statement	Exemplary quote
Reduced brand visibility	Brands will have reduced visibility on voice assistants compared to other touchpoints.	<i>"During a product search, by the time you get to the third item, you have forgotten what the first was and what the price of the second one was. You're done beyond the third results."</i> Dr. A. K. Pradeep, CEO at MachineVantage
Rising private label strength	Alexa will disproportionately place its private labels while penalizing other consumer brands.	<i>"If I ask Alexa to send me twenty AA batteries, I will probably get Amazon's branded batteries. However, if I explicitly ask for Duracell, I receive my preferred brand, provided it is available on the platform. Thus, companies have to invest in branding even more than they did before so that consumers asked for a product by the name"</i> Jim Sterne, Emeritus Director of the Digital Analytics Association
Increasing cost of paid media	The cost of advertising on voice assistants will be higher than web-based advertising because of the limited space available for sponsored messages (paid recommendations).	<i>"Voice commerce brings up a "real estate" problem. While I can display several ads on the same Google Search results page, I don't have the same ad space on smart speakers. Thus, I expect the cost of voice ads to be more than two times higher than regular search ads. Am I able to justify this cost increase?"</i> Maurizio Miggiano, Head of Digital at Generali
Reduced access to consumer data	Brands will not have access to consumer insights related to shopper's behavior and interest as this information will not be shared by the voice assistant's manufacturer.	<i>"Amazon observes every consumer's step and uses this information to customize campaigns. My problem is not necessarily to pay Alexa or similar platforms to reach my consumers but that I do not have individual-level data needed to develop the relationship with consumers further"</i> Lorenzo Farronato - VP Marketing Communications at Swarovski
Universality of the impact across categories	Low involvement product categories will not be the only ones affected by the voice assistant's diffusion.	<i>"It doesn't matter whether you are selling gold jewellery or a Rolls-Royce. Alexa does commoditize entire product categories, all the way from diamonds to detergents. Your brand has become a commodity fighting for air space"</i> Dr. A. K. Pradeep, CEO at MachineVantage
Increasing competition due to constant recommendation assessment	Voice assistants will on-goingly re-evaluate the consumer's product choice and suggest better alternatives. In doing so, brands are required to justify their positions constantly.	<i>"When consumers will ask Alexa to repurchase a preferred brand of whisky, for instance, the voice assistant will answer "I have discovered a new brand of whisky which is proven to be healthier, cheaper and appreciated by other users. Would you like to try it?". In this context, brands will constantly need to justify their leadership position in the consumer's shopping cart"</i> Cosimo Accoto, Research Affiliate at MIT

Table 1. Effect of the diffusion of voice commerce on brand dynamics

Although all six anticipated consequences for brands have been consistently mentioned by the informants, three of them seem to have a high business impact for top management. Thus, the potential reduction of brand visibility via organic search results, rise of retailers' private labels, and increase in media spending are discussed in detail.

Search algorithms represent the gatekeeper for modern companies and retailers. As reported by informants, compared to display-enabled devices, the optimization of voice search results on VAs presents three structural challenges due to the nature of consumer interactions and information framing. First, during voice shopping users can review one to three options before they start forgetting information such as price or quantity of the mentioned products. Reduced attention span and short-term memory can negatively influence the satisfaction toward this shopping system, especially when the user is required to search for products in an explorative way extensively. Second, VAs deliver search results to users in the form of recommendations. The assistive nature of the interaction with VAs implies a delegation of responsibility, at least in the absence of explicit requests by the user. Whenever a user directly asks for a specific brand or product, VAs respond with the closest option available to them. However, when shopping for items without specifying a brand, VAs are more likely to recommend their private labels, if available. In the case of Alexa, when a brand name is not proactively mentioned by the user, the private label, under the name of Amazon's Choice, appears as the first recommendation in over 50% of instances (Cheris et al., 2017). Third, the search engine results continuously adapt to the user's purchase history and the evol-

ing understanding that VA acquires about its interlocutor. However, after a user has purchased a product, for example, Nespresso coffee capsules, the subsequent suggestions for coffee start from the same manufacturer. As such, this dynamic might reduce variety seeking in shoppers.

Private label development is seen as particularly dangerous by national brands (see Quelch & Harding, 1996). In utilitarian product categories characterized by low purchase involvement, the parallel expansion of private labels and VAs represent a risk for category 'commoditization' (Sterne, 2017). An emblematic example of this process comes from the battery business. A few years after its launch in 2009, Amazon's private label 'AmazonBasics' accounts for 31% of the overall battery sales online by large margins from national brands such as Duracell (21%) and Energizer (12%) (Neff, 2016). With Amazon's private label portfolio growing to 135 brands and more than 330 Amazon exclusive brands, similar trends gradually become visible in a variety of product categories such as skincare, home improvement tools, and golf equipment. The limited 'shelf space' available to merchants on in-home smart devices strengthens the private brands' position. According to Cheris et al. (2017), for categories in which Amazon offers private-label products, Alexa recommends the private-label products 17% of the time, although these products represent only about 2% of the total volume sold. Amazon's biased placement on VAs of its private labels against national brands challenges the traditional retail marketing practice that expects a distribution of a given brand, 'share of shelf', proportional to its 'market share'. Furthermore, consumers can decide to automate fully (e.g., subscription) or semi-automate

(e.g., product added to the shopping list) their purchases creating self-established lock-in mechanisms.

Advertising represented for decades the primary tool to generate brand awareness, improving both recall and recognition. With the rise of the Internet, the concept of advertising transmuted to search engines where advertisers buy promotional spaces in response to a set of keywords searched by the user. Within digital advertising, 'search advertising' represents the most successful format, accounting for 45% of the total spending (IAB & PWC, 2018). Advertisers face an overall cost increase of search ads with a particular impact on highly competitive consumer products. For instance, the cost per click on the search term 'laundry detergent liquid' reached \$17 on Amazon in a given period (Cheris et al., 2017). Search advertising in the form of voice has a paramount role in voice commerce marketing. Although brands are generally positive toward this new form of investment, the peculiarities of the voice channel induce concerns. Compared to web browser navigation, where search engines can display ten results per page and up to five advertisements, VAs can only suggest a few results with limited space for sponsored messages. This scarcity of space might increase competition among advertisers with a consequent rise in advertising costs.

Managerial Implications and Conclusions

As voice assistants become better at learning consumer preferences and habits, they introduce consumer biases and heuristics likely to affect marketing and e-commerce practices. The exponential development of

these AI-enabled devices is altering shopping behavior and posing new challenges and opportunities for managers. Key informants predict that VAs will assume a central relational role in the consumer market and progressively mediate market interactions. Simultaneously, they expect VA's intrinsic social nature to increasingly influence consumer decision-making while the relationship with this new media shifts to steadfast dependency. These fast-changing market dynamics within the context of voice shopping may have a severe impact on consumer brands. Loss of brand visibility, the increased relevance of retailers' private labels, and the growth in advertising costs are the three main anticipated consequences for brands among the six identified in our study. In light of these potential dynamics, managers need to explore the challenges and opportunities posed by voice technologies and prepare to respond to this disruption in the media ecosystem. Future research may employ an ecosystem perspective to study more closely the adoption of AI-based VAs by companies as well as the interplay between machine, brand, and consumer behavior (Rahwan et al., 2019). Furthermore, media management and innovation experts are encouraged to explore how the introduction of VAs may alter the media landscape and impact the management and digital transformation of legacy media organizations.

References

- Abdollahpouri, H., Adomavicius, G., Burke, R., Guy, I., Jannach, D., Kamishima, T., Krasnodebski, J., & Pizzato, L. (2020). Beyond Personalization: Research Directions in Multistakeholder Rec-

- ommendation. *User Modeling and User-Adapted Interaction*, 30(1), 127–158. <https://doi.org/10.1007/s11257-019-09256-1>
- Ammari, T., Kaye, J., Tsai, J. Y., & Bentley, F. (2019). Music, Search, and IoT: How People (Really) Use Voice Assistants. *ACM Transactions on Computer-Human Interaction*, 26(3), 17:1–17:28. <https://doi.org/10.1145/3311956>
- André, Q., Carmon, Z., Wertenbroch, K., Crum, A., Frank, D., Goldstein, W., Huber, J., Van Boven, L., Weber, B. and Yang, H. (2018). Consumer choice and autonomy in the age of artificial intelligence and big data. *Customer Needs and Solutions*, 5(1-2), pp.28–37. <https://doi.org/10.1007/s40547-017-0085-8>
- Araujo, T. (2018). Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Computers in Human Behavior*, 85, 183–189. <https://doi.org/10.1016/j.chb.2018.03.051>
- Aronovitch, C. D. (1976). The Voice of Personality: Stereotyped Judgments and their Relation to Voice Quality and Sex of Speaker. *The Journal of Social Psychology*, 99(2), 207–220. <https://doi.org/10.1080/00224545.1976.9924774>
- Avis, M. (2003). Do we need methodological theory to do qualitative research?. *Qualitative health research*, 13(7), 995–1004.
- Chérif, E., & Lemoine, J.-F. (2019). Anthropomorphic virtual assistants and the reactions of Internet users: An experiment on the assistant's voice. *Recherche et Applications En Marketing (English Edition)*, 34(1), 28–47. <https://doi.org/10.1177/2051570719829432>
- Cheris, A., Rigby, D., & Tager, S. (2017). *Dreaming of an Amazon Christmas?* Bain & Company. Available at: <https://www.bain.com/insights/retail-holiday-newsletter-2017-issue-2> (Accessed: 3 August 2020).
- Colleoni, E., Rozza, A., & Arvidsson, A. (2014). Echo Chamber or Public Sphere? Predicting Political Orientation and Measuring Political Homophily in Twitter Using Big Data. *Journal of Communication*, 64(2), 317–332. <https://doi.org/10.1111/jcom.12084>
- Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative sociology*, 13(1), 3–21.
- Davenport, T., Guha, A., Grewal, D., & Bressgott, T. (2020). How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science*, 48(1), 24–42. <https://doi.org/10.1007/s11747-019-00696-0>
- Dawar, N., & Bendle, N. (2018). Marketing in the age of Alexa. *Harvard Business Review*, 96(3), 80–86.
- Dellaert, B. G. C., Shu, S. B., Arentze, T. A., Baker, T., Diehl, K., Donkers, B., Fast, N. J., Häubl, G., Johnson, H., Karmarkar, U. R., Oppewal, H., Schmitt, B. H., Schroeder, J., Spiller, S. A., & Steffel, M. (2020). Consumer decisions with artificially intelligent voice assistants. *Marketing Letters*, 31(4), 335–347. <https://doi.org/10.1007/s11002-020-09537-5>
- eMarketer (2019). *Which Select Activities Have US Smart Speaker Owners Done on Their Smart Speakers*. Available at: <https://www.emarketer.com/chart/230338/> (Accessed: 3 April 2019).
- Gao, F. (2019). *Use New Alexa Emotions and Speaking Styles to Create a More Natural and Intuitive Voice Experience*. Available at: <https://developer.amazon.com/en-US/blogs/alexa> (Accessed: 10 July 2020).
- Gao, F. (2020). *Create Engaging Voice Experiences with a New Alexa Speaking Style, Speaking Styles in Amazon Polly Voices and Additional Polly Voices*. Available at: <https://developer>

- amazon.com/en-US/blogs/alexa (Accessed: 10 July 2020).
- Gefen, D., & Straub, D. W. (2004). Consumer trust in B2C e-Commerce and the importance of social presence: Experiments in e-Products and e-Services. *Omega*, 32(6), 407–424. <https://doi.org/10.1016/j.omega.2004.01.006>
- Giles, H. (1970). Evaluative Reactions to Accents. *Educational Review*, 22(3), 211–227. <https://doi.org/10.1080/0013191700220301>
- Glaser, B. G., & Strauss, A. L. (2017). *Discovery of grounded theory: Strategies for qualitative research*. Routledge.
- Go, E., & Sundar, S. S. (2019). Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior*, 97, 304–316. <https://doi.org/10.1016/j.chb.2019.01.020>
- Han, S., & Yang, H. (2018). Understanding adoption of intelligent personal assistants: A parasocial relationship perspective. *Industrial Management & Data Systems*, 118(3), 618–636. <https://doi.org/10.1108/IMDS-05-2017-0214>
- Hansen, J. H. L. (1996). Analysis and compensation of speech under stress and noise for environmental robustness in speech recognition. *Speech Communication*, 20(1), 151–173. [https://doi.org/10.1016/S0167-6393\(96\)00050-7](https://doi.org/10.1016/S0167-6393(96)00050-7)
- Holzwarth, M., Janiszewski, C., & Neumann, M. M. (2006). The Influence of Avatars on Online Consumer Shopping Behavior. *Journal of Marketing*, 70(4), 19–36. <https://doi.org/10.1509/jmkg.70.4.019>
- Hoy, M. B. (2018). Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants. *Medical Reference Services Quarterly*, 37(1), 81–88. <https://doi.org/10.1080/02763869.2018.1404391>
- Huang, M.-H., & Rust, R. T. (2018). Artificial Intelligence in Service. *Journal of Service Research*, 21(2), 155–172. <https://doi.org/10.1177/1094670517752459>
- IAB & PWC (2019). *IAB Internet Advertising Revenue Report 2018 Full Year Results*. Available at: <https://www.iab.com/insights/iab-internet-advertising-revenue-report-2018-full-year-results/> (Accessed: 10 July 2020).
- Ishowo-Oloko, F., Bonnefon, J.F., Soroye, Z., Crandall, J., Rahwan, I. and Rahwan, T. (2019). Behavioural evidence for a transparency–efficiency tradeoff in human–machine cooperation. *Nature Machine Intelligence*, 1(11), pp.517–521. <https://doi.org/10.1038/s42256-019-0113-5>
- Jeffrey, R. C. (1974). Preference among preferences. *The Journal of Philosophy*, 71(13), 377–391.
- Knote, R., Janson, A., Söllner, M., & Leimeister, J. M. (2019). *Classifying Smart Personal Assistants: An Empirical Cluster Analysis*. Hawaii International Conference on System Sciences (HICSS), Maui, Hawaii, USA. <https://www.alexandria.unisg.ch/255501/>
- Kumar, V., Dixit, A., Javalgi, R. (Raj) G., & Dass, M. (2016). Research framework, strategies, and applications of intelligent agent technologies (IATs) in marketing. *Journal of the Academy of Marketing Science*, 44(1), 24–45. <https://doi.org/10.1007/s11747-015-0426-9>
- Labecki, A. J., Klaus, P., & Zaichkowsky, J. L. (2018). *How Bots Have Taken over Brand Choice Decisions*. https://doi.org/10.1007/978-3-030-02683-7_72
- Lee, S., & Choi, J. (2017). Enhancing user experience with conversational agent for movie recommendation. *International Journal of Human-Computer Studies*, 103(C), 95–105. <https://doi.org/10.1016/j.ijhcs.2017.02.005>

- Lester, J. C., Converse, S. A., Kahler, S. E., Barlow, S. T., Stone, B. A., & Bhogal, R. S. (1997). The persona effect: Affective impact of animated pedagogical agents. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, 359–366. <https://doi.org/10.1145/258549.258797>
- Lu, B., Fan, W., & Zhou, M. (2016). Social presence, trust, and social commerce purchase intention: An empirical research. *Computers in Human Behavior*, 56, 225–237. <https://doi.org/10.1016/j.chb.2015.11.057>
- Mandelli, A. (2018). *Intelligenza artificiale e marketing. Agenti invisibili, esperienza, valore e business*. EGEA.
- Mari, A., & Algesheimer, R. (2021a). *The Role of Trusting Beliefs in Voice Assistants during Voice Shopping* (p. 4073). <https://doi.org/10.24251/HICSS.2021.495>
- Mari, A., & Algesheimer, R. (2021b). AI-based Voice Assistants for Digital Marketing: Preparing for Voice Marketing and Commerce. In *Contemporary issues in Digital Marketing*. Routledge.
- Mari, A., Mandelli, A., & Algesheimer, R. (2020). The Evolution of Marketing in the Context of Voice Commerce: A Managerial Perspective. In F. F.-H. Nah & K. Siau (Eds.), *HCI in Business, Government and Organizations* (pp. 405–425). Springer International Publishing. https://doi.org/10.1007/978-3-030-50341-3_32
- McLean, G., Osei-Frimpong, K., & Barhorst, J. (2021). Alexa, do voice assistants influence consumer brand engagement? – Examining the role of AI powered voice assistants in influencing consumer brand engagement. *Journal of Business Research*, 124, 312–328. <https://doi.org/10.1016/j.jbusres.2020.11.045>
- Nass, C. I., & Brave, S. (2005). *Wired for speech: How voice activates and advances the human-computer relationship* (p. 9). Cambridge, MA: MIT press.
- Nass, C., & Gong, L. (2000). Speech interfaces from an evolutionary perspective. *Communications of the ACM*, 43(9), 36–43. <https://doi.org/10.1145/348941.348976>
- Nass, C., & Moon, Y. (2000). Machines and Mindlessness: Social Responses to Computers. *Journal of Social Issues*, 56(1), 81–103. <https://doi.org/10.1111/0022-4537.00153>
- Nass, C., & Steuer, J. (1993). Voices, Boxes, and Sources of Messages: Computers and Social Actors. *Human Communication Research*, 19(4), 504–527. <https://doi.org/10.1111/j.1468-2958.1993.tb00311.x>
- Neff, J. (2016). *Amazon's private labels already dominate battery and speaker sales online*. AdAge. Available at: <https://adage.com/article/digital/amazon-private-label-dominates-batteries-speakers/306602> (Accessed: 13 July 2020).
- Norvig, S. R. (2011). *Artificial Intelligence: A Modern Approach, Global Edition* (3rd edition). Pearson.
- Olshavsky, R. W., & Granbois, D. H. (1979). Consumer Decision Making-Fact or Fiction? *Journal of Consumer Research*, 6(2), 93–100.
- Qiu, L., & Benbasat, I. (2009). Evaluating Anthropomorphic Product Recommendation Agents: A Social Relationship Perspective to Designing Information Systems. *Journal of Management Information Systems*, 25(4), 145–182. <https://doi.org/10.2753/MIS0742-1222250405>
- Quelch, J. A., & Harding, D. (1996). Brands versus private labels: Fighting to win. *Harvard Business Review*, 74(1), 99–109.
- Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J.-F., Breazeal, C., Crandall, J. W., Christakis, N. A., Couzin, I. D., Jackson, M. O., Jennings, N. R., Kamar, E., Kloumann, I. M., Larochelle, H., Lazer, D., McElreath, R., Mislove, A., Parkes, D. C., Pentland, A. "Sandy," ... Well-

- man, M. (2019). Machine behaviour. *Nature*, 568(7753), 477–486. <https://doi.org/10.1038/s41586-019-1138-y>
- Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people*. Cambridge, UK: Cambridge university press.
- Rhee, C. E., & Choi, J. (2020). Effects of personalization and social role in voice shopping: An experimental study on product recommendation by a conversational voice agent. *Computers in Human Behavior*, 109, 106359. <https://doi.org/10.1016/j.chb.2020.106359>
- Ram, A., Prasad, R., Khatri, C., Venkatesh, A., Gabriel, R., Liu, Q., Nunn, J., Hedayatnia, B., Cheng, M., Nagar, A. and King, E. (2018). Conversational ai: The science behind the alexa prize. arXiv preprint arXiv:1801.03604.
- Sarikaya, R. (2017). The Technology Behind Personal Digital Assistants: An overview of the system architecture and key components. *IEEE Signal Processing Magazine*, 34(1), 67–81. <https://doi.org/10.1109/MSP.2016.2617341>
- Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication*, 40(1), 227–256. [https://doi.org/10.1016/S0167-6393\(02\)00084-5](https://doi.org/10.1016/S0167-6393(02)00084-5)
- Schmitt, B. (2019). From Atoms to Bits and Back: A Research Curation on Digital Technology and Agenda for Future Research. *Journal of Consumer Research*, 46(4), 825–832. <https://doi.org/10.1093/jcr/ucz038>
- Shen, A. (2014). Recommendations as personalized marketing: Insights from customer experiences. *Journal of Services Marketing*, 28(5), 414–427. <https://doi.org/10.1108/JSM-04-2013-0083>
- Short, R. V. (1976). Definition of the problem-The evolution of human reproduction. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 195(1118), 3–24.
- Solarino, A. M., & Aguinis, H. (2020). Challenges and best practice recommendations for designing and conducting interviews with elite informants. *Journal of Management Studies*.
- Sterne, J. (2017). *Artificial intelligence for marketing: practical applications*. John Wiley & Sons. ISBN: 978- 1-119-40633-4
- Von der Puetten, A.M., Krämer, N.C., Gratch, J. and Kang, S.H. (2010). "It doesn't matter what you are!" explaining social effects of agents and avatars. *Computers in Human Behavior*. <https://doi.org/10.1016/j.chb.2010.06.012>
- Voosen, P. (2017). The AI detectives. *Science*, 357(6346), 22–27. <https://doi.org/10.1126/science.357.6346.22>
- Wang, L. C., Baker, J., Wagner, J. A., & Wakefield, K. (2007). Can A Retail Web Site be Social? *Journal of Marketing*, 71(3), 143–157. <https://doi.org/10.1509/jmkg.71.3.143>

About the Author

Alex Mari

Alex Mari is a Research Associate and PhD candidate for the Chair for Marketing and Market Research, University of Zurich. He is visiting PhD student at the Säid Business School, University of Oxford. Alex studies the adoption of AI-based voice assistants for marketing initiatives and their effect on shopping behavior. He advises entrepreneurs and teaches digital and AI marketing in business schools. Previously, Alex worked as a director of digital marketing (Sonova), brand manager (P&G) and tech startup CEO. He holds an MSc in Marketing from the University of Lugano.

ORCID ID: <https://orcid.org/0000-0003-4803-5096>