"I Will Follow You!" – How Recommendation Modality Impacts Processing Fluency and Purchase Intention

Completed Research Paper

Melanie Schwede
University of Goettingen
Goettingen, Germany
melanie.schwede@uni-goettingen.de

Naim Zierau
University of St.Gallen
St. Gallen, Switzerland
naim.zierau@unisg.ch

Andreas Janson
University of St.Gallen
St. Gallen, Switzerland
andreas.janson@unisg.ch

Maik Hammerschmidt
University of Goettingen
Goettingen, Germany
maik.hammerschmidt@wiwi.uni-goettingen.de

Jan Marco Leimeister
University of St.Gallen
St. Gallen, Switzerland
janmarco.leimeister@unisg.ch

Abstract

Although conversational agents (CA) are increasingly used for providing purchase recommendations, important design questions remain. Across two experiments we examine with a novel fluency mechanism how recommendation modality (speech vs. text) shapes recommendation evaluation (persuasiveness and risk), the intention to follow the recommendation, and how modality interacts with the style of recommendation explanation (verbal vs. numerical). Findings provide robust evidence that text-based CAs outperform speech-based CAs in terms of processing fluency and consumer responses. They show that numerical explanations increase processing fluency and purchase intention of both recommendation modalities. The results underline the importance of processing fluency for the decision to follow a recommendation and highlight that processing fluency can be actively shaped through design decisions in terms of implementing the right modality and aligning it with the optimal explanation style. For practice, we offer actionable implications on how to make effective sales agents out of CAs.

Keywords: Conversational Agent, Recommendation Modality, Explanation of Recommendation, Processing Fluency, Purchase Intention

Introduction

Firms increasingly use conversational agents (CA) based on artificial intelligence (AI) for providing purchase recommendations in order to facilitate consumers' decision-making (Chernev 2003) and boost sales (Karnik 2019). For example, retailers such as Sephora employ CAs to provide virtual product consultancy and shopping assistance (Arthur 2016). CAs represent a novel interaction paradigm between consumers and firms and have been coined the "next operating system in commerce" (Tricks 2021). They allow firms to engage consumers in an immediate, dialogue-based interaction to provide information on products and services (Miao et al. 2022). The provision of recommendations by CAs can occur either through speech or text (Wirtz et al. 2018). Consumers are increasingly interested in purchasing products via speech-based CAs, e.g., for food (Schwartz 2019), resulting in a 120% increase in the number of "speech commerce" users in the last three years (Kinsella 2021). Given these modality options (speech- vs. text-based), it is crucial for firms to understand how the modality of purchase recommendations influences the chance that consumers follow the recommendation and purchase a product. However, research on how different recommendation modalities impact downstream consequences for consumers' experience (information processing and evaluation of recommendation persuasiveness and adoption risk) and firms in turn (purchase), and how recommendations should be "designed" for different modalities has been scarce. So far, modality-related research in the technology context has focused on the influence of modality on preference expression (Klesse et al. 2015), product search (King et al. 2021), user flow experience (Zierau et al. 2022), the facilitation of different consumer tasks (Rzepka et al. 2021) and communication styles (Whang and Im 2021) as well as the influence of the (in)congruence between speaking and listening on recommendation acceptance (Hu et al. 2022).

We have long known from communication theories (i.e., media richness, media synchronicity) that speech- and text-based modalities differ fundamentally in their characteristics (Schmitt et al. 2021a). Specifically, speech is more synchronous and provides a richer symbol set than text. Therefore, various verbal cues such as pitch or volume are part of the symbol set of speech and can change the understanding of a message within a certain time (Moffett et al. 2021). It is therefore likely that the same message is processed differently when read than when heard. For instance, in the advertising context, it has been shown that communication modalities provide different metacognitive experiences, ranging from fluency to disfluency (Fransen et al. 2010). Thus, certain modalities of providing a purchase recommendation are more likely to be processed smoothly than others. In recommendation situations with CAs, where consumers are exposed to a relatively novel and uncertain decision environment, whether recommendations are easily accessible can make a difference for the purchase decision. It is therefore essential for firms to understand how modality influences the processing of purchase recommendations so that consumers feel well advised in their product choice and in turn purchase the recommended product.

However, for getting consumers to follow a purchase recommendation, a crucial issue is that they understand why a CA provided the recommendations (Cheng and Jiang 2022). A widely proposed mechanism for creating transparent purchase recommendations is to provide explanations for the recommendations (Schmitt et al. 2022; Wang and Benbasat 2007). Explanations have been shown to support consumers to effectively use CAs (Wang and Benbasat 2007) and to make better product decisions (Gregor and Benbasat 1999). In the context of e-commerce, explanations can be communicated in a numerical or verbal style. Numerical explanations include information that contains specific numbers, e.g., "87% of consumers purchased this product" while verbal explanations consist of information that includes general descriptions, e.g., "many of the consumers purchased this product". As different explanation styles are processed and remembered differently (Viswanathan and Childers 1996), it is likely that certain combinations of recommendation modality (speech vs. text) and explanation style are more easily processed than others. To date, no study has examined whether aligning recommendation modality and recommendation explanation style makes a difference for purchase intentions.

Taken together, this paper aims to address the following two research questions:

RQ1: *How does the modality for providing recommendations through conversational agents influence consumers' interaction experience (processing fluency, recommendation persuasiveness, and recommendation adoption risk) and purchase intention?*

RQ2: *How does the interplay of recommendation modality and explanation style influence consumers' interaction experience and purchase intention?*

To answer these research questions, we conducted two experimental studies in which we assessed participants' experience when interacting with a speech- vs. text-based CA to receive a food product recommendation. The findings of this research highlight a novel explanatory mechanism that impacts how recommendation modality affects consumer decision-making across different explanation styles. These findings have important research implications for a theory-driven design of conversational systems as well as practical implications for firms that aim to leverage CAs as a novel sales channel.

The structure of the paper is as follows: Initially, we explain processing fluency theory as the conceptual basis of this study. Then, we introduce the state of research on the two key constructs recommendation modality and explanation of recommendation. Subsequently, we present the research model and develop the hypotheses. Our research model was tested in two empirical studies. For this purpose, we outline the study design, samples, and measures of the two studies. The results are next presented and discussed. We complete the paper with theoretical and practical contributions, limitations, and future research directions.

Theoretical Background and Related Work

The following section provides the theoretical background for our research. Therefore, we discuss the relevance of processing fluency for facilitating interaction evaluation and purchase intention in the context of interactions with CA and elaborate on how processing fluency is shaped by design stimuli (modality and explanation of recommendation). Further, we present related work on recommendation modalities and explanations of recommendations in the service context.

The Role of Processing Fluency in Purchase Contexts

Communication properties, such as speech vs. text, are related to whether consumers perceive an interaction with a CA as easy or difficult to process (Song and Schwarz 2009; Zierau et al. 2022). Processing fluency describes the subjective feelings of ease or difficulty that consumers experience when mentally processing information, which is also referred to as meta-cognitive experience (Graf et al. 2018; Novemsky et al. 2007). According to Graf et al. (2018), every mental operation such as perceiving, processing, and retrieving information are related to the consumer's feeling of fluency. Moreover, processing fluency is a central construct that influences consumer evaluation of a (decision-making) situation as likable or risky (Graf et al. 2018) as well as consumer behavior across a wide array of domains (Schwarz et al. 2021). Accordingly, if consumers find it easier to process information via speech or text, then this has a direct impact on their decisions.

A concept somewhat related to fluency is cognitive effort, i.e., the perceived time and effort required to complete a specific task (Le Bigot et al. 2007; Mosteller et al. 2014; Rzepka et al. 2021). However, the consequences of cognitive effort for consumer responses depend on the level of fluency of information processing associated with these cognitive investments (Mosteller et al. 2014). Hence, a high cognitive effort for a task needs not to be detrimental for consumer responses if it entails high fluency of information processing. Therefore, we argue that the level of processing fluency underlying a CA interaction is a more decisive mechanism for explaining the effects of CAs on consumer responses. In line with this argument, extant research on how cognitive effort associated with speech- vs. text-based CAs impacts consumers' subsequent responses (Le Bigot et al. 2007; Rzepka et al. 2021) provides inconclusive findings, likely because the degrees of processing fluency underlying the interaction with the examined CAs and the resulting cognitive effort varied across these studies. No study so far has examined how different modalities of CA interactions influence processing fluency (Schwarz et al. 2021). We address this research gap by examining the impact of the modality of CAs, specifically relating to the provision of recommendations, on processing fluency.

Further, we argue that two key recommendation evaluations strongly influenced by processing fluency are recommendation persuasiveness and recommendation adoption risk. Recommendation persuasiveness describes the positive and convincing feeling of a consumer towards the recommendations of a CA (Lehto et al. 2012; Shevechuk et al. 2019). Only when consumers are persuaded by a recommendation through a CA, they are willing to purchase a product (Cialdini 2009; Rhee and Choi 2020). While consumers may be persuaded of the purchase recommendation's fit with their needs, they may still feel a certain risk of adopting the recommendation, either socially or financially. The recommendation adoption risk is defined as the consumers' evaluation “of uncertainty and potentially adverse consequences of buying a recommended product” (Xiao and Benbasat 2007, p. 145). Consumers evaluate the anticipated risk of a failure of a recommendation in advance, i.e., before they purchase the product or service. That is, consumers are uncertain whether a product will work as expected, which can lead to a fear of financial loss or that the recommended product will be disapproved by a group of peers (Lee et al. 2022; Song and Schwarz 2009; Tsiros and Heilman 2005). The risk evaluations in turn may prevent consumers from buying a recommended product (Tsiros and Heilman 2005).

As already mentioned, the implementation and design of explanations for a recommendation are relevant in the context of different CAs. Previous research on explanations has focused on online consumer ratings presented in a mean vs. distribution format (Kostyk et al. 2017). However, no study so far investigated whether giving an explanation has any effect at all on the processing fluency of recommendations and subsequent recommendation evaluations (persuasiveness and risk) and whether there is a difference if it is communicated in a numerical or verbal style.

Related Work on Recommendation Modalities

Until now, most studies on CAs have focused on the design and acceptance of text-based agents (Araujo 2018; Hill et al. 2015; Huang and Rust 2018), but the emergence of speech-based interactions in commercial activities creates new challenges for the effective presentation of purchase recommendations. Firms are particularly interested in determining whether and how to implement speech-based recommendations (Bentahar 2018). Specifically, the communication-related affordances of speech-based interactions differ fundamentally to text-based modalities that are likely to differentially affect consumer experience of the purchase recommendation process.

According to Moffett et al. (2021), modalities differ regarding the degree of synchronicity. Speech-based recommendation modalities are characterized by a high degree of synchronicity as speech-based communication creates an immediate interaction scenario, in which consumers are involved in a constant back and forth. In a speech-based interaction, consumers must adapt to the speed of the conversation partner. In addition, consumers cannot revise and refine a message; speech-based interactions require an immediate, unfiltered response (Schmitt et al. 2021b). In contrast, text-based interactions typically allow for a delayed communication and thus have a low degree of synchronicity. In text communication, consumers can read and understand another's message at their own speed and compose their response in the same way (Le Bigot et al. 2007). Furthermore, modalities differ in terms of the richness of symbols. Speech-based interactions provide a richer symbol set than text-based interactions. Specifically, speech-based interactions allow for the integration of vocal features (i.e., changes in frequency and amplitude) and prosody (i.e., pauses). Therefore, according to media richness theory, speech-based interactions are “richer” than text-based interactions and afford consumers to send a variety of social cues making communication more socially enjoyable (Rzepka et al. 2021).

Despite these fundamental differences between speech and text modality, there is limited research on how speech-based interactions determine the effectiveness of recommendations. The emerging literature on speech-based recommendation has focused on mechanisms related to anthropomorphism and its downstream consequences for consumers and firms, showing that speech-based recommendations can enhance social experiences in e-commerce settings (Hess et al. 2009; Qiu and Benbasat 2005, 2009). However, research on how speech-based CAs affect information processing is scarce. Initial research has shown that speech-based preference expression can lead to more indulgent product choice when ordering a product with a vending machine (Klesse et al. 2015), hinting at modality-induced differences in consumer decision-making processes. Looking at search modalities, King et al. (2021) show that speech-based searches decrease purchase intention based on a more deliberative mindset. Similarly, taking a task-technology-fit perspective, Rzepka et al. (2021) could show that depending on the task type speech- vs. text-based CAs differentially affect the level of effort (and enjoyment) associated with a service encounter when receiving a restaurant recommendation. The initial studies underline the importance of processing fluency as a potent psychological mechanism to explain the effectiveness of message design across contexts. Building on this work, we aim to investigate how the modality of a recommendation impacts the cognitive processing of a purchase recommendation.

Related Work on Explanations of Recommendations

CAs are able to provide customized recommendations based on the needs and desires of consumers articulated during an interaction. Firms increasingly provide recommendations in product-related CA interactions to use such agents not only for providing information and services but to sell products (Karnik

2019). For instance, they often recommend an appropriate product to consumers from a wide range of choices and seek to ease consumers' workload and facilitate purchase (Aksoy et al. 2006).

In the context of CAs, a recommendation is based either on collaborative filtering, i.e., products are recommended based on purchases of similar consumers, or on content-based filtering, i.e., products are recommended based on the consumer's preferences and needs (Gai and Klesse 2019). Firms often use a hybrid form of this filtering for their algorithm. Current research, therefore, focuses on the impact of the recommendation algorithm (collaborative- vs. content-based vs. hybrid). However, consumers can only understand how the CA algorithm works by explaining how a recommendation was made, irrespective of the type of algorithm. Research has shown that an explanation about the provided recommendation leads to higher trust in CAs (Wang and Benbasat 2007).

Typically, firms use two fundamentally different communication styles when explaining product features: numerical and verbal communication. However, these are two different heuristics that consumers process differently (Viswanathan and Childers 1996). Numerical information like “83% of consumers also bought...” encompasses specific attributes while verbal information like “most of the consumers also bought...” comprises a generic description of attributes (Viswanathan and Childers 1996). However, up to now, no research exists on which explanation style is most effective in terms of enhancing the fluency of processing a recommendation. Moreover, it is unknown so far whether a certain explanation style fits a certain modality better than others. We contribute to this gap and examine which recommendation explanation style should be used for a specific modality.

Hypotheses Development

In this section, we first introduce the research model based on the previous fluency theory as well as the previous insights in literature and develop the single hypotheses of the research model. The hypotheses development follows in three steps. Initially, we infer the effect of recommendation modality on processing fluency and in turn on recommendation evaluation (recommendation persuasiveness and recommendation adoption risk) and purchase intention. These hypotheses are tested in study 1. In a second step, we hypothesize the effect of explanation of recommendation on consumer responses, while in a third step, we elaborate the effects of the interaction between recommendation modality and explanation of recommendation. These effects regarding the second and third steps are tested in the second study.

Research Model

In order to understand what impact our three stimuli (recommendation modality, recommendation explanation, and their interaction) have on processing fluency of the recommendation and what influence this central mechanism has on the subsequent evaluation of the recommendation, and finally on behavioral response in terms of purchase intention, we established the following research model, which is presented in Figure 1. To analyze the influence of the recommendation modality and recommendation explanation as well as their interaction, they were added as independent variables to the research model. The processing fluency in turn mediates the effect of the three stimuli on the recommendation persuasiveness and the recommendation adoption risk, which in turn influence purchase intention. Hence, consumers' psychological processes capture the direct processing of the stimuli in a fluent or disfluent manner and the resulting evaluation of perceived persuasiveness and risk of the provided recommendation.

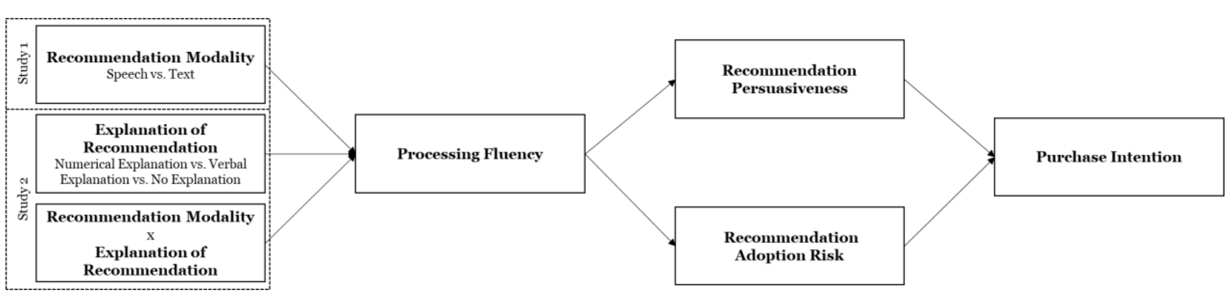


Figure 1. Research Model

The Effect of Recommendation Modality on Experience and Purchase Intention

CAs who provide purchase recommendations typically reduce the amount of information by minimizing the selection of recommended products as well as sharing only the most important details of the products. As previously indicated, speech conveys a higher degree of richness in comparison to text and is therefore advantageous when exchanging ambiguous or social information. However, richness is not efficient for exchanging unambiguous information, such as recommendations, so text is more targeted and therefore easier to process (Moffett et al. 2021). Verbal information in contrast to text are temporary and have to be hold in mind, resulting in a higher load on working memory and a reduced processing quality (Baddeley and Hitch 1974). Therefore, condensing information increases complexity and mental workload within speech-based interaction as it is more difficult to listen than to read (Hong et al. 2004; Le Bigot et al. 2007). Using speech to communicate seems more natural but when consumers have to compare (purchase) recommendations, especially in e-commerce, consumers usually get a visual information in form of text or pictures for the reduction of complexity. As a result, consumers have become accustomed to perceiving and processing visual (text) information at their own speed compared to auditory (speech) information, making it easier for them to retrieve information.

In addition, modalities are associated with different presentation styles of information (simultaneous vs. sequential) which can also increase or decrease complexity (Basu and Savani 2019; Mogilner et al. 2013). Simultaneous presentation of information is a feature of text-based interaction that eases information processing (Schmitt et al. 2021b) as it lowers the difficulty of comparison (Basu and Savani 2019). In contrast, consumers find it more difficult to assimilate and remember information via speech (Le Bigot et al. 2007), especially when comparing between options. When CAs provide purchase recommendations, a comparison of at least two product options occurs. Therefore, the text modality seems to fit better as it presents information simultaneously and allows the consumer to compare product options. Thus, we obtain the following hypothesis:

H1: Consumers perceive the processing of recommendations by a speech-based conversational agent (vs. a text-based conversational agent) as less fluent.

Processing fluency is described as a subjective feeling that leads consumers to make evaluations and show behaviors in response to this good or bad feeling of fluency or disfluency (Schwarz et al. 2021). Research findings indicate that stimuli that are perceived as easy to process are also perceived as more positive by consumers (Winkielman et al. 2003). The positive feeling leads to a more reliable evaluation of information and a certain ease within the decision process which increases the general liking of certain information (Mosteller et al. 2014; Schwarz et al. 2021). As a result, consumers are likely to be persuaded by the recommendation they have received. Hence, a positive feeling can be seen as a heuristic that consumers use to evaluate the persuasiveness of a situation (Chen et al. 1999). To ensure that consumers consider a recommendation as persuasive, recommendations need to convey arguments or cues that influence consumer evaluations and responses (Cialdini 2009; Dehnert and Mongeau 2022). The positive feeling of high persuasiveness gives the impression that a recommendation is beneficial for the consumer (Kahneman and Frederick 2007). In other words, it is convincing and leads to a certain behavior, in the case of purchase recommendations, to buy a recommended product. This leads us to the following two hypotheses:

H2a: If processing fluency is high, recommendation persuasiveness increases.

H2b: If recommendation persuasiveness is high, purchase intention increases.

Furthermore, a fluent information processing can contribute to a feeling of familiarity and thus reduce the perceived risk of a hazardous situation or decision (Schwarz et al. 2021). Thus, the ease of processing information can lead to a perceived sense of security (Dohle and Montoya 2017). On the contrary, if consumers perceive information as disfluent, this has a negative impact on risk evaluation (Song and Schwarz 2009). When consumers receive recommendations from a new technology, such as CAs, consumers have a lower level of initial trust towards the technology in comparison to humans (Xiao and Benbasat 2007). However, if the recommendation modality can cause a feeling of processing fluency and thus a familiar impression, then this also has a positive effect on the consumer's risk evaluation, in form of feeling less uncertain. If consumers judge the adoption of a recommendation as less risky, they are more likely to purchase a product (Kahneman and Frederick 2007). In line with this argument, previous research has shown that consumers are less willing to buy products when they are in a rather risky decision context (Xiao and Benbasat 2007). We, therefore, hypothesize the following:

H3a: If processing fluency is high, recommendation adoption risk decreases.

H3b: If recommendation adoption risk is high, purchase intention decreases.

The Effect of Explanation of Recommendations on Processing Fluency

The communication of an explanation for recommendations can be considered as a cue to help consumers to make decisions. This type of stimuli reveals the CA's reasoning line, which is based on consumer needs and product feature preferences, thereby outlining the logical processes of the recommendations. Therefore, any type of explanation closes the knowledge gap between the CA and consumer, leading to a positive feeling towards the reliability of the CA and its recommendations (Wang and Benbasat 2007). The processing fluency theory indicates that in particular the feeling of reliability, which an explanation for the recommendation conveys, leads to an increase of processing fluency (Song and Schwarz 2009).

Firms typically use the collaborative approach for explanations due to the benefits in the search context of products (Liao and Sundar 2021). Moreover, collaborative explanations fuel the bandwagon heuristic. As Sundar (2008) outlines, consumers like to go with the flow of people and assume that if a large number of their fellows agree that something is good, then consumers will be more likely to adopt that mindset. This kind of heuristic leads consumers to be able to make decisions without thinking much about it (Kahneman and Frederick 2007). Therefore, consumers do not have to question the recommendations at all and have to pay less attention, which leads to an increased processing fluency of the recommendation. Consequently, an (numerical or verbal) explanation of a recommendation provides consumers a feeling of reliance. Moreover, the collaborative explanation triggers social validation in consumers, as they would like to act in line with others. These heuristics enable consumers to process information more easily (Benner et al. 2021). Thus, we obtain the following hypothesis:

H4: A (a) numerical and (b) verbal explanation of the recommendations (vs. no explanation) increases processing fluency.

Due to the effects of processing fluency on perceived recommendation persuasiveness, recommendation adoption risk and purchase intention hypothesized in the previous section, and the assumption that explanations have a positive effect on processing fluency, explanations of the recommendations also have a positive effect on subsequent evaluations and outcome variables.

The Interplay of Recommendation Modality and Explanation of Recommendation

As mentioned above, in collaborative explanations for a purchase recommendation, firms relate to a reference group (i.e., other consumers) which is either expressed with specific numbers like “XX% of consumers similar to you” or descriptive words like “most consumers similar to you” (Liao and Sundar 2021). However, numerical information often do not have inherent meaning, whereas verbal information can more effortlessly convey meaning (Viswanathan and Childers 1996). Therefore, numerical information require greater cognitive effort, making these information easier to be overlooked, particularly comparing this to vivid or verbal information (Zillmann and Brosius 2012), which can convey meaning more easily. However, presenting numerical information in text-based form can perform better than verbal information for product comparisons (Viswanathan and Childers 1996). The higher degree of accurate encoding in comparison tasks offsets the disadvantage of transforming numerical information to extract its inherent meaning. Therefore, in text-based interactions, more complex (i.e., numerical) information can be processed, while speech-based interaction is better suited for conveying simpler (i.e., verbal) information (Schmitt et al. 2021b). Accordingly, we assume that a numerical explanation is considered easier to process in interactions with text-based CAs and a verbal explanation is easier to process with speech-based CAs, thus the following hypotheses emerge:

H5a: A numerical explanation (vs. verbal explanation) of the recommendations by a text-based conversational agent increases processing fluency.

H5b: A verbal explanation (vs. numerical explanation) of the recommendations by a speech-based conversational agent increases processing fluency.

Experimental Procedures and Research Context

To test our research model, we used a custom-made conversational interface building on *Python* and *Google WaveNet* (Zierau et al. 2022). The interface provides the ability to model both speech- and text-based conversational interactions. Across modalities, the design of the interface was kept simple to control for potential confounds and to isolate the effect of recommendation modality on consumer experience. For both studies, the dialogue flow is predefined to ensure that the content and structure of the conversational interaction is kept stable across participants. For the speech-based interface, we used a state-of-the-art text-to-speech generator from Google WaveNet to create the different voice prompts (van den Oord et al. 2016).

As a study context, we chose purchase recommendations within the food industry for three major reasons. First, the food sector was one of the first contexts in which purchase recommendations have been made via a speech-based CA. Especially when cooking, the advantages of a speech-based CA can come into play in terms of hands-free operation (De Bellis and Johar 2020; Sciuto et al.). Second, consumers can readily imagine food products and the level of product complexity is low. Thus, it is not essential that the consumer needs a picture to evaluate the usefulness of a product. Third, in recent years, numerous skills for speech- and text-based CAs have been developed to provide consumers with recipe tips and recommendations for food products. Thus, the scenario exhibits a high degree of realism, which enhances the external validity of our studies.

In this paper, we present evidence from two studies that were designed to test our hypotheses.

Study 1

Study 1 was designed to provide a first test of our baseline hypotheses that speech-based (as opposed to text-based) interactions evoke less processing fluency when getting a purchase recommendation. Furthermore, study 1 also tests whether increases in processing fluency increase recommendation persuasiveness and decrease the risk of adopting recommendations, and whether these changes increase purchase intention.

Design and Procedure

We recruited a representative US consumer sample using the crowdsourcing platform Prolific. Our sample consists of 181 participants which had to speak English as a primary language and needed a working microphone and loudspeaker. The participants received financial compensation after completing the online study. In addition, we followed the guidelines of Lowry et al. (2016) for recruiting participants on crowdsourcing platforms (e.g., using attention checks, submission rate). Those who failed to answer attention or treatment checks correctly were disregarded from further analyses. Thus, the effective sample resulted in a total of 141 participants (M_{Age}=34.59, SD_{Age}=13.26, 48.23% female). Participants were randomly assigned to a two-cell between subject experiment (speech- vs. text-based interaction). In both conditions, the experimental task comprised getting a purchase recommendation on peanut butter. Before the interaction with the conversational agent started, we asked participants to emerge themselves in a scenario describing an event, where they were preparing a desert for their friends. In the scenario, we provided them with a set of preferences related to product features (e.g., maximum price), and they were instructed to report to a fictitious app called *KitchenHelper* (via speech or text depending on the condition). We asked participants to report all of these preferences to receive a purchase recommendation. The sequence and semantics of the questions the interface was posing to participants were identical across conditions and participants (i.e., the sequence of eight conversational turns, semantics, and all other features of the purchase recommendation process). Moreover, to ensure comparability, in the text-based condition each prompt was displayed individually. The participants were only allowed to use laptops or desktop computers to avoid confounding effects based on device.

Immediately after the experimental task, we assessed the constructs of the research model. Table 1 shows all items of the constructs, the item loadings as well as the Cronbach's alphas; all of them are above the cut-off value of 0.7, indicating construct-level reliability. We also captured a battery of control measures including measures on social demographics, scenario realism, and personal traits and experiences with CAs.

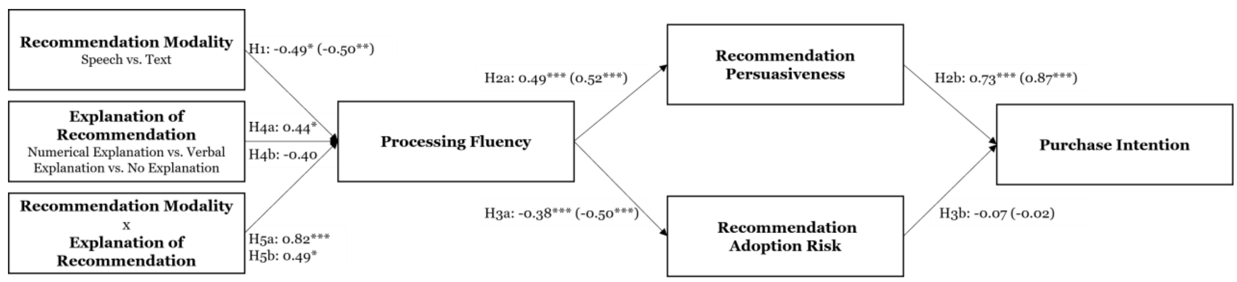
Constructs	Items	Item Loadings	
		Study 1	Study 2
Processing Fluency adapted from Graf et al. (2018) Study 1: $\alpha = 0.91$ Study 2: $\alpha = 0.91$	The process of studying the food product recommendations of the agent was …		
	(1) difficult / (7) easy	0.90	0.90
	(1) disfluent / (7) fluent	0.83	0.87
	(1) incomprehensible / (7) comprehensible	0.85	0.89
	(1) unclear / (7) clear	0.89	0.88
	(1) effortful / (7) effortless	0.83	0.74
Recommendation Persuasiveness adapted from Shevechuk et al. (2019) Study 1: $\alpha = 0.81$ Study 2: $\alpha = 0.82$	The food product recommendations of the agent had an influence on me.	0.88	0.88
	The food product recommendations of the agent were personally relevant for me.	0.87	0.87
	In my opinion, the food product recommendations of the agent were convincing.	0.81	0.83
Recommendation Adoption Risk adapted from Song and Schwarz (2009) Study 1: $\alpha = 0.93$ Study 2: $\alpha = 0.93$	How did you rate the risk of adopting a food product recommendation from the agent?		
	(1) very safe / (7) very risky	0.92	0.92
	(1) not very hazardous / (7) very hazardous	0.93	0.96
	(1) very harmless / (7) very harmful	0.96	0.95
Purchase Intention adapted from Bleier et al. (2019) and King et al. (2021) Study 1: $\alpha = 0.87$ Study 2: $\alpha = 0.85$	It is likely that I would purchase at least one of the recommended food products I searched for.	0.94	0.93
	I am ready to purchase at least one of the recommended food products I searched for.	0.94	0.91
	It is not probable that I would purchase at least one of the recommended food products I searched for. (R)	0.80	0.80

Notes: α = Cronbach's alpha; (R) *reversed item*

Table 1. Measures of Constructs, Indicator, and Construct Reliability

Results

To test H_1, we conducted a one-way analysis of variance (ANOVA). The ANOVA results show a significant main effect of recommendation modality on processing fluency ($F(1,140) = 7.87$, $p < 0.01$). A planned contrast of predictive margins reveals that participants who used the speech-based CA experienced significantly lower levels of processing fluency compared to participants using the text-based CA ($M_{\text{Speech}} = 5.93$, $SE = 0.12$; $M_{\text{Text}} = 6.43$, $SE = 0.13$; $t = -2.81$, $p < 0.01$). Thus, we can confirm H_1. Furthermore, we conducted a seemingly unrelated regression (SUR). Since the dependent variables of the processing fluency model are simultaneously the independent variables of the purchase intention model, it seems unrealistic for them to have independent error terms because they are related in content. Therefore, the method of SUR is able to provide more efficient estimates for coefficients (Zellner 1962). The path model results are summarized in Figure 2. These findings demonstrate that the decrease of processing fluency evoked by the speech-based CA ($\beta_{\text{Speech}} = -0.50$, $SE = 0.18$, $p < 0.01$) in turn led to a significant decrease in recommendation persuasiveness ($\gamma_{\text{Fluency}} = 0.52$, $SE = 0.09$, $p < 0.001$) and a significant increase in recommendation adoption risk ($\delta_{\text{Fluency}} = -0.50$, $SE = 0.10$, $p < 0.001$). Focusing on the impact on purchase intention, we found that only the negative impact of speech-based CA on recommendation persuasiveness, in turn, led to a significant decrease in purchase intention ($\zeta_{\text{Persuasiveness}} = 0.87$, $SE = 0.10$, $p < 0.001$), while the increase in recommendation adoption risk had no significant impact on purchase intention ($\zeta_{Risk} = 0.02$, $SE = 0.09$, $p = 0.86$). Thus, we were able to confirm hypotheses H_{2a}, H_{2b}, and H_{3a} and had to reject hypothesis H_{3b}.



Notes: *p ⩽ 0.05; **p ⩽ 0.01; ***p ⩽ 0.001; text and values in parentheses refer to study 1

Figure 2. Path Model of Study 1 & 2

Discussion

The findings of study 1 provide initial evidence for our baseline hypothesis that a speech-based CA creates less processing fluency compared to text-based CA. Moreover, we can show that these differences lead to reduced recommendation persuasiveness and increased risk evaluation when using a speech-based CA. However, only recommendation persuasiveness can be linked to purchase intention.

Study 2

The main objective of study 2 was to examine to which extent the presence of an explanation for the recommendation can either enhance or reduce participants' experience of processing fluency. Moreover, we distinguished between explanations that either used a numerical or a verbal style.

Design and Procedure

Again, we recruited a total of 362 participants from the crowdsourcing platform Prolific. We used the previously described criteria of the recruitment again for study 2. We disregarded those participants who either failed to answer our attention checks correctly or those who did not pass the treatment checks. Specifically, those who did not recognize whether the explanation used a numerical vs. verbal style were disregarded from the sample. The final sample consisted of 293 participants (M_{Age}=32.62, SD_{Age}=12.63, 56.66% female). Participants were randomly assigned to a 2 (recommendation modality: speech-based vs. text-based) x 3 (explanation: numerical vs. verbal vs. no explanation) between-subjects design, using the same experimental paradigm and context as in study 1 and only adding a short explanation. To distinguish between the different explanation conditions, we kept the content and meaning stable and only shifted between a numerical and verbal communication style. Specifically, the numerical condition used numbers (i.e., "*Based on preferences of 91.8% of the Kitchen Helper App users and about 652,000 reviews of around 17 peanut butter products*, […]"), the verbal condition used words (i.e., "*Based on preferences from most of the Kitchen Helper App users and a lot of reviews of many peanut butter products*, […]"). For the no explanation condition, we added a short sentence stating that the CA had considered all available options, but did not offer any explanation. To control for the length of the turn and to isolate the effect of the explanation, we held the number of words constant between conditions. Immediately, after the experimental task we collected the same measures as in study 1 (see Table 1). All constructs and items show reliable values.

Results

To test H_1, H_4, and H_5, we conducted a one-way analysis of variance (ANOVA). The results of ANOVA show a significant main effect of recommendation modality on processing fluency ($F(1, 287) = 8.31$, $p < 0.01$). Planned contrasts of predictive margins show that speech is significantly less fluent than text ($M_{Speech} = 5.74$, *SE* = 0.09; M_{Text} = 6.12, *SE* = 0.09; *t* = –2.88, $p < 0.01$), which provides support for H_1. Also, the main effect of providing an explanation on processing fluency was significant ($F(2, 287) = 8.90$, $p < 0.001$). Again, planned contrasts of predictive margins show that numerical explanation is significantly more fluent than no explanation ($M_{NumericalExplanation}$ = 6.29, *SE* = 0.11; $M_{NoExplanation}$ = 5.86, *SE* = 0.12; *t* = 2.73, $p < 0.01$). But there is no significant difference between verbal explanation and no explanation ($M_{VerbalExplanation}$ = 5.64, *SE* = 0.12; *t* = -1.35, *p* = 0.180). Therefore, we could confirm H_{4a} and had to reject H_{4b}. We use planned contrasts to analyze the interaction effect of recommendation modality and explanation of recommendation. Figure 3 shows the predictive margins of the interaction effects. We find support for H_{5a}, as the processing fluency is higher for numerical explanations than for verbal explanations provided by text-based CAs ($M_{Numerical \times Text}$ = 6.54, *SE* = 0.16; $M_{Verbal \times Text}$ = 5.72, *SE* = 0.16; *t* = 3.59, $p < 0.001$). However, we find no support for H_{5b}. Instead, we identified a reversed effect, as the numerical explanation provided by speech-based CAs has a significantly higher effect on fluency than the verbal explanation ($M_{Numerical \times Speech}$ = 6.05, *SE* = 0.14; $M_{Verbal \times Speech}$ = 5.56, *SE* = 0.17; *t* = 2.21, $p < 0.05$).

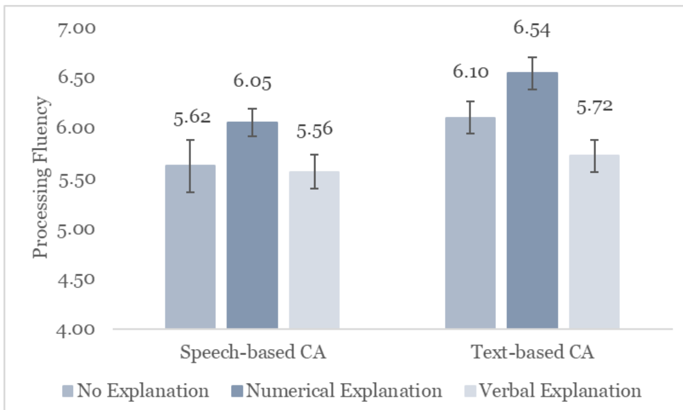


Figure 3. Predictive Margins of the Interaction Effects for Processing Fluency

Next, we estimated a path model using the SUR function in STATA to assess the overall system of hypotheses, see Figure 2. These findings demonstrate that the decrease of processing fluency evoked by the speech-based CA (β_{Speech} = -0.49, *SE* = 0.23, $p < 0.05$) in turn led to a significant decrease in consumers' evaluation of recommendation persuasiveness ($\gamma_{Fluency}$ = 0.49, *SE* = 0.06, $p < 0.001$) and a significant increase in consumers' evaluation of recommendation adoption risk ($\delta_{Fluency}$ = -0.38, *SE* = 0.08, $p < 0.001$). Focusing on the impact on purchasing intention, we found that only the negative impact of speech-based CA on recommendation persuasiveness, in turn, led to a significant decrease in purchase intention ($\zeta_{Persuasiveness}$ = 0.73, *SE* = 0.06, $p < 0.001$), while the increase in recommendation risk had no significant impact on purchase intention (ζ_{Risk} = -0.07, *SE* = 0.05, *p* = 0.15). The results of the SUR also show that a numerical explanation (vs. no explanation; $\beta_{NumericalExplanation}$ = 0.44, *SE* = 0.22, $p < 0.05$) increases the processing fluency significantly, but a verbal explanation (vs. no explanation) had no effect ($\beta_{VerbalExplanation}$ = -0.40, *SE* = 0.22, *p* = 0.1). As in study 1, we were able to confirm hypotheses H_{2a}, H_{2b}, and H_{3a} and had to reject hypothesis H_{3b}. Moreover, supporting H_{5a}, a positive significant interaction effect of the text-based CA and the numerical vs. verbal explanation of recommendations on processing fluency is shown ($\beta_{Numerical \times Text}$ = 0.82, *SE* = 0.23, $p < 0.001$). Finally, rejecting $H_{5b,}$ as the numerical vs. verbal explanation of recommendations provided by speech-based CAs has a significantly positive effect on processing fluency ($\beta_{Numerical \times Speech}$ = 0.49, *SE* = 0.22, $p < 0.05$).

Discussion

The results of study 2 provide further evidence for our initial hypothesis that speech-based interactions produce lower processing fluency compared to text-based interactions. In addition, we repeatedly demonstrate that processing fluency reduces the risk evaluation of recommendation adoption and increases recommendation persuasiveness. Recommendation persuasiveness also increases purchase intention, whereas risk evaluation does not, as in study 1. Thus, processing fluency is an important mechanism that determines whether consumers purchase a recommended product. Moreover, numerical explanation of recommendations leads to higher processing fluency of the recommendation for both speech- and text-based CAs.

General Discussion

Firms face the challenge to decide which type of CAs they want to use for communicating purchase recommendations because it is well established that CAs can help consumers to make better product decisions (Gai and Klesse 2019). We aimed to show how fluently purchase recommendations can be processed via speech- and text-based CAs and what effects this has on the evaluation of the recommendations and the corresponding purchase intentions. In addition, we aimed to show if an explanation of the recommendations can contribute to the ease of processing purchase recommendations. Since speech- vs. text-based interactions comprise fundamental differences (Schmitt et al. 2021a), and yet explanations are a central component of recommendations, it is even more relevant to understand whether and how explanations have an effect on processing fluency within different recommendation modalities.

We show in both studies that purchase recommendations provided by text-based CAs are perceived as easier to process as compared to speech-based CAs. Speech-based interactions are frequently hyped as a new technology due to their naturalness and flexibility, leading to a growing interest in integrating them into the purchasing context (Schwartz 2022). However, shopping, especially for new products, is more complex in nature, as consumers want to feel they are being well advised by receiving customized recommendations as well as understanding why they are receiving certain products from a CA (Gai and Klesse 2019; Wang and Benbasat 2007). Across two studies, our results demonstrate that even when only two purchase recommendations need to be compared, the text-based CA outperforms the speech-based CA. Moreover, our research results indicate that processing fluency and its theoretical approach is a good explanatory mechanism to show that recommendations are processed differently via speech- vs. text-based CAs. Thus, we enrich and extend cognitive effort research within the modality context, which comes to different conclusions (Le Bigot et al. 2007; Rzepka et al. 2021). Therefore, further research should include the construct of processing fluency as a predictor of cognitive effort (i.e., the cognitive effort it takes to complete a task) and its effects on service outcome.

The fluency of processing purchase recommendations is a central mechanism, which primarily conveys the feeling of whether recommendations can be processed as easy or difficult. This feeling in turn leads to an evaluation of the recommendation and the intention to follow the purchase recommendation by purchasing the product(s). Our results show that processing fluency, which is associated with feelings of ease, significantly minimized perceived recommendation adoption risk and increased recommendation persuasiveness. We also found that perceived persuasiveness significantly increased the intention to purchase a recommended product. However, the recommendation adoption risk did not significantly influence the purchase intention. A possible explanation for the non-significant effect is the food product used in our experiment. Food products such as peanut butter are low-involvement products for which the risk of purchasing the recommended product is manageable. However, in the context of shopping with CAs, low-involvement products are currently the standard.

The second study examined a further stimulus on processing fluency, the explanation for the recommendations. Furthermore, we examined whether there was an interaction effect between the type of CA and the explanation for recommendations. A numerical explanation was significantly perceived as more fluent than no explanation. However, a verbal explanation was not perceived as more fluent than no explanation, i.e., there were no significant differences between the two styles. Thus, we note that a numerical explanation is more specific than a verbal explanation, which is more generically in nature (Viswanathan and Childers 1996). Not necessarily processing ability, but processing preferences may make a difference in whether and which style of explanation is preferred by the consumer (Childers et al. 1985). Since consumers are confronted with numbers regularly when making product decisions and since these numbers are more unambiguous and easier to interpret, a numerical explanation is suitable for processing purchase recommendations more fluently. Interestingly, numerical explanations (as opposed to verbal explanations) have a significantly higher influence on processing fluency for both speech- and text-based CAs. In particular, numbers can be identified as a heuristic and a kind of anchor. Processing numbers in speech-based interactions requires some cognitive effort, yet these numbers do not need to be internalized in their entirety (Viswanathan and Childers 1996), and still convey a sense of accuracy when the numbers are highly marked (Liao and Sundar 2021; Schmitt et al. 2021b). The feeling of seeing or hearing a high number conveys the feeling of an easier processing of the purchase recommendations because the recommendations can be better categorized.

Our two studies indicate that consumers process the purchase recommendation of a speech-based CA less fluently compared to a text-based CA. Firms should therefore carefully consider prioritizing a text-based CA for digital commerce over a speech-based CA or consider how they can increase the fluency of the speech-based CA. The use of a numerical explanation can already increase the fluency of speech-based CAs, but also of text-based CAs, so that they are preferable in everyday commerce contexts to ultimately increase the purchase intention. Given the increase of the number of available CAs in the omnichannel landscape, such a generalizable conversational design is important to provide superior consumer experiences across modalities.

Implications, Limitations, and Future Research

Our study has several important implications for IS theory and practice. First, we provide a rigorous experimental approach to disentangle the effects of CA modality as one key material property of these IT artifacts. Across two studies we highlight that in shopping and purchase recommendation contexts, text-based CAs usually outperform speech-based CAs. Consequently, we challenge the assumption that the naturalness of speech is superior compared to text-based interactions. Second, and well connected to our first theory implication, with our fluency theory perspective we contribute to a better understanding of CA interaction. By highlighting the crucial role of fluent interaction on consumer evaluation, we underscore the impact that fluency theory could have in understanding consumer reactions to CAs. Third, we highlight that not only modality and fluency matter for CA interaction outcomes. Research should also very carefully consider the content-wise perspective when highlighting crucial aspects of interactions, i.e., in our case recommendation explanations. By showing that numerical explanations outperform no explanations and verbal explanations, we identify important anchoring points during conversational interactions.

Consequently, we highlight as a major practical implication that firms should use text-based CAs in shopping contexts to communicate purchase recommendations and enrich them with a numerical explanation for the recommendations, so that the overall processing fluency is high, leading to the persuasion that the recommended product is appropriate and hence making the consumer more likely to purchase the product.

As advances in technology enable firms to implement both speech- and text-based CAs, and consumers' requests for these technologies increase, firms should consider how they will design each modality. Therefore, it is also important to take a multi- and omnichannel perspective. In the long term, it will not be a question of whether firms use speech- or text-based CAs, but rather consumers will decide which CA they prefer and when, so both CAs will become equally relevant. However, our research has highlighted potential problems with speech-based CAs. Moreover, our findings show that even for simple products that do not require a complicated explanation, consumers' processing fluency is significantly reduced with speech-based interfaces. Due to the synchronous nature of speech-based interactions, it is to be expected that this deleterious effect on processing fluency is amplified for product combinations that require careful deliberation (e.g., comparing multiple product combinations). At the same time, prior research has shown that speech-based interfaces can positively affect a wide range of affective user experiences including social presence (Qiu and Benbasat 2009) and flow (Zierau et al. 2022). As such, future research should take a contextual perspective and investigate when using speech-based CAs becomes beneficial or even more detrimental in e-commerce contexts. Moreover, researchers could assess conversational designs that enhance processing fluency to build on the potential of speech-based interfaces to boost affective user evaluations and, ultimately, product purchases.

Our paper has naturally limitations that provide even more ground for future research. We conducted our two studies online by utilizing an online panel. Thus, we acknowledge that we could not control for all external stimuli in comparison to in-person lab experiments. Such noise could especially influence our central fluency variable. However, our setting is comparable to typical shopping experiences and, thus, we took this limitation consciously into account. Field experiments could thus take our findings as a starting point to provide evidence if our theory holds in large-scale shopping settings. To avoid any co-effects that may threaten internal validity, we used an instantiation of a CA that was reduced to its core interactive features. Hence, future work may assess the effect of recommendation modality in field settings to reliably infer effect sizes against the background of more complex CA designs.

In the experiment, participants were asked to put themselves in a service encounter with the CA that was as realistic as possible. More specifically, this meant that they had to put themselves in a specific role (vignette), act accordingly, and make decisions. This type of experiment has limitations in external validity (Aguinis and Bradley 2014). Therefore, the study's results have limited generalizability to real-life situations in service encounters with CAs. Moreover, the experiment depicts a short-term and fixed process. Future experiments should also investigate more flexible service situations (along the lines of Wizard-of-Oz experiments) and their impact. Furthermore, future studies should test different products and situations in the context of purchase recommendations to strengthen the robustness of our results and discuss possible further effects.

When explaining a recommendation, we chose a collaborative explanation type in the experiment, however, there is also the variation of a content-based explanation, which should be further explored in future studies. Moreover, we designed the collaborative explanation using numerical and verbal elements. However, future studies should use alternative numerical and verbal elements to validate the results of our study. The sequence of recommendations, the presented product features, and social communication elements also comprise a recommendation and should be investigated in the context of speech- vs. text-based CAs.

Acknowledgements

We thank the Swiss National Science Foundation for funding parts of this research (100013_192718). The third author also acknowledges funding from the Basic Research Fund (GFF) of the University of St. Gallen.

References

Aguinis, H., and Bradley, K. J. 2014. "Best Practice Recommendations for Designing and Implementing Experimental Vignette Methodology Studies," *Organizational Research Methods* (17:4), pp. 351-371 (doi: 10.1177/1094428114547952).

Aksoy, L., Bloom, P. N., Lurie, N. H., and Cooil, B. 2006. "Should Recommendation Agents Think Like People?" *Journal of Service Research* (8:4), pp. 297-315 (doi: 10.1177/1094670506286326).

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), pp. 183-189 (doi: 10.1016/j.chb.2018.03.051).

Arthur, R. 2016. "Sephora Launches Chatbot On Messaging App Kik," available at https://www.forbes.com/sites/rachelarthur/2016/03/30/sephora-launches-chatbot-on-messaging-app-kik/?sh=72577169547e, accessed on May 1 2022.

Baddeley, A. D., and Hitch, G. 1974. "Working Memory," Elsevier, pp. 47-89 (doi: 10.1016/S0079-7421(08)60452-1).

Basu, S., and Savani, K. 2019. "Choosing Among Options Presented Sequentially or Simultaneously," *Current Directions in Psychological Science* (28:1), pp. 97-101 (doi: 10.1177/0963721418806646).

Benner, D., Schöbel, S., and Janson, A. 2021. "Exploring the State-of-the-Art of Persuasive Design for Smart Personal Assistants," in *Proceedings of the 16th International Conference on Wirtschaftsinformatik*, Duisburg-Essen, Germany.

Bentahar, A. 2018. "How voice shopping is changing shopping," available at https://www.forbes.com/sites/forbesagencycouncil/2018/12/03/how-voice-search-is-changing-shopping/?sh=177940026152, accessed on Feb 14 2022.

Bleier, A., Harmeling, C. M., and Palmatier, R. W. 2019. "Creating Effective Online Customer Experiences," *Journal of Marketing* (83:2), pp. 98-119 (doi: 10.1177/0022242918809930).

Chen, S., Duckworth, K., and Chaiken, S. 1999. "Motivated Heuristic and Systematic Processing," *Psychological Inquiry* (10:1), pp. 44-49 (doi: 10.2307/1449522).

Cheng, Y., and Jiang, H. 2022. "Customer–brand relationship in the era of artificial intelligence: understanding the role of chatbot marketing efforts," *Journal of Product & Brand Management* (31:2), pp. 252-264 (doi: 10.1108/JPBM-05-2020-2907).

Chernev, A. 2003. "When More Is Less and Less Is More: The Role of Ideal Point Availability and Assortment in Consumer Choice," *Journal of Consumer Research* (30:2), pp. 170-183 (doi: 10.1086/376808).

Childers, T. L., Houston, M. J., and Heckler, S. E. 1985. "Measurement of Individual Differences in Visual versus Verbal Information Processing," *Journal of Consumer Research* (12:2), pp. 125-133 (doi: 10.1086/208501).

Cialdini, R. B. 2009. *Influence*: *Science and practice*, Boston, Mass.: Pearson Education.

De Bellis, E., and Johar, G. V. 2020. "Autonomous Shopping Systems: Identifying and Overcoming Barriers to Consumer Adoption," *Journal of Retailing* (96:1), pp. 74-87 (doi: 10.1016/j.jretai.2019.12.004).

Dehnert, M., and Mongeau, P. A. 2022. "Persuasion in the Age of Artificial Intelligence (AI): Theories and Complications of AI-Based Persuasion," *Human Communication Research* (48:3), pp. 386-403 (doi: 10.1093/hcr/hqac006).

Dohle, S., and Montoya, A. K. 2017. "The dark side of fluency: Fluent names increase drug dosing," *Journal of Experimental Psychology* (23:3), pp. 231-239 (doi: 10.1037/xap0000131).

Fransen, M. L., Fennis, B. M., and Pruyn, A. T. H. 2010. "Matching Communication Modalities: The Effects of Modality Congruence and Processing Style on Brand Evaluation and Brand Choice," *Communication Research* (37:4), pp. 576-598 (doi: 10.1177/0093650210368251).

Gai, P. J., and Klesse, A.-K. 2019. "Making Recommendations More Effective Through Framings: Impacts of User- Versus Item-Based Framings on Recommendation Click-Throughs," *Journal of Marketing* (83:6), pp. 61-75 (doi: 10.1177/0022242919873901).

Graf, L. K. M., Mayer, S., and Landwehr, J. R. 2018. "Measuring Processing Fluency: One versus Five Items," *Journal of Consumer Psychology* (28:3), pp. 393-411 (doi: 10.1002/jcpy.1021).

Gregor, S., and Benbasat, I. 1999. "Explanations from Intelligent Systems: Theoretical Foundations and Implications for Practice," *MIS Quarterly* (23:4), pp. 497-530 (doi: 10.2307/249487).

Hess, T., Fuller, M., and Campbell, D. 2009. "Designing Interfaces with Social Presence: Using Vividness and Extraversion to Create Social Recommendation Agents," *Journal of the Association for Information Systems* (10:12), pp. 889-919 (doi: 10.17705/1jais.00216).

Hill, J., Ford, W. R., and Farreras, I. G. 2015. "Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations," *Computers in Human Behavior* (49:3), pp. 245-250 (doi: 10.1016/j.chb.2015.02.026).

Hong, W., Thong, J. Y., and Tam, K. Y. 2004. "The Effects of Information Format and Shopping Task on Consumers' Online Shopping Behavior: A Cognitive Fit Perspective," *Journal of Management Information Systems* (21:3), pp. 149-184 (doi: 10.1080/07421222.2004.11045812).

Hu, P., Gong, Y., Lu, Y., and Ding, A. W. 2022. "Speaking vs. listening? Balance conversation attributes of voice assistants for better voice marketing," *International Journal of Research in Marketing* (doi: 10.1016/j.ijresmar.2022.04.006).

Huang, M.-H., and Rust, R. T. 2018. "Artificial Intelligence in Service," *Journal of Service Research* (21:2), pp. 155-172 (doi: 10.1177/1094670517752459).

Kahneman, D., and Frederick, S. 2007. "Frames and brains: elicitation and control of response tendencies," *Trends in cognitive sciences* (11:2), pp. 45-46 (doi: 10.1016/j.tics.2006.11.007).

Karnik, A. 2019. "The Who, When And Why Of Delivering Personalized Product Recommendations," available at https://www.forbes.com/sites/forbescommunicationscouncil/2019/04/22/the-who-when-and-why-of-delivering-personalized-product-recommendations/?sh=87a0f3e1b925, accessed on May 1 2022.

King, D., Auschaitrakul, S., and Lin, C.-W. J. 2021. "Search modality effects: merely changing product search modality alters purchase intentions," *Journal of the Academy of Marketing Science*, 1-12 (doi: 10.1007/s11747-021-00820-z).

Kinsella, B. 2021. "Voice Shopping Rises to 45 Million U.S. Adults in 2021," available at https://voicebot.ai/2021/12/24/voice-shopping-rises-to-45-million-u-s-adults-in-2021/#:~:text=In%202018%2C%20Voicebot%20Research%20found,Shopping%20Consumer%20Adoption%20Report%202021., accessed on May 1 2022.

Klesse, A.-K., Levav, J., and Goukens, C. 2015. "The Effect of Preference Expression Modality on Self-Control," *Journal of Consumer Research* (42:4), 535–550 (doi: 10.1093/jcr/ucv043).

Kostyk, A., Niculescu, M., and Leonhardt, J. M. 2017. "Less is more: Online consumer ratings' format affects purchase intentions and processing," *Journal of Consumer Behaviour* (16:5), pp. 434-441 (doi: 10.1002/cb.1643).

Le Bigot, L., Rouet, J.-F., and Jamet, E. 2007. "Effects of speech- and text-based interaction modes in natural language human-computer dialogue," *Human factors* (49:6), pp. 1045-1053 (doi: 10.1518/001872007X249901).

Lee, S., Oh, J., and Moon, W.-K. 2022. "Adopting Voice Assistants in Online Shopping: Examining the Role of Social Presence, Performance Risk, and Machine Heuristic," *International Journal of Human–Computer Interaction*, pp. 1-15 (doi: 10.1080/10447318.2022.2089813).

Lehto, T., Oinas-Kukkonen, H., and Drozd, F. 2012. "Factors Affecting Perceived Persuasiveness of a Behavior Change Support System," in *Proccedings of the 33rd International Conference on Information Systems*, Orlando, USA.

Liao, M., and Sundar, S. S. 2021. "When E-Commerce Personalization Systems Show and Tell: Investigating the Relative Persuasive Appeal of Content-Based versus Collaborative Filtering," *Journal of Advertising*, pp. 1-16 (doi: 10.1080/00913367.2021.1887013).

Lowry, P. B., D'Arcy, J., Hammer, B., and Moody, G. D. 2016. ""Cargo Cult" science in traditional organization and information systems survey research: A case for using nontraditional methods of data collection, including Mechanical Turk and online panels," *The Journal of Strategic Information Systems* (25:3), pp. 232-240 (doi: 10.1016/j.jsis.2016.06.002).

Miao, F., Kozlenkova, I. V., Wang, H., Xie, T., and Palmatier, R. W. 2022. "An Emerging Theory of Avatar Marketing," *Journal of Marketing* (86:1), pp. 67-90 (doi: 10.1177/0022242921996646).

Moffett, J. W., Folse, J. A. G., and Palmatier, R. W. 2021. "A theory of multiformat communication: mechanisms, dynamics, and strategies," *Journal of the Academy of Marketing Science* (49:3), pp. 441-461 (doi: 10.1007/s11747-020-00750-2).

Mogilner, C., Shiv, B., and Iyengar, S. S. 2013. "Eternal Quest for the Best: Sequential (vs. Simultaneous) Option Presentation Undermines Choice Commitment," *Journal of Consumer Research* (39:6), pp. 1300-1312 (doi: 10.1086/668534).

Mosteller, J., Donthu, N., and Eroglu, S. 2014. "The fluent online shopping experience," *Journal of Business Research* (67:11), pp. 2486-2493 (doi: 10.1016/j.jbusres.2014.03.009).

Novemsky, N., Dhar, R., Schwarz, N., and Simonson, I. 2007. "Preference Fluency in Choice," *Journal of Marketing Research* (44:3), pp. 347-356 (doi: 10.1509/jmkr.44.3.347).

Qiu, L., and Benbasat, I. 2005. "Online Consumer Trust and Live Help Interfaces: The Effects of Text-to-Speech Voice and Three-Dimensional Avatars," *International Journal of Human-Computer Interaction* (19:1), pp. 75-94 (doi: 10.1207/s15327590ijhc1901_6).

Qiu, L., and Benbasat, I. 2009. "Evaluating Anthropomorphic Product Recommendation Agents: A Social Relationship Perspective to Designing Information Systems," *Journal of Management Information Systems* (25:4), pp. 145-182 (doi: 10.2753/MIS0742-1222250405).

Rhee, C. E., and 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* (109:1), p. 106359 (doi: 10.1016/j.chb.2020.106359).

Rzepka, C., Berger, B., and Hess, T. 2021. "Voice Assistant vs. Chatbot – Examining the Fit Between Conversational Agents' Interaction Modalities and Information Search Tasks," *Information Systems Frontiers* (24), pp. 839-856 (doi: 10.1007/s10796-021-10226-5).

Schmitt, A., Wambsganss, T., and Janson, A. 2022. "Designing for Conversational System Trustworthiness: The Impact of Model Transparency on Trust and Task Performance," in *Proccedings of the 30th European Conference on Information Systems*, Timişoara, Romania.

Schmitt, A., Wambsganss, T., Söllner, M., and Janson, A. 2021a. "Exploring the Gap Between Perceived Trust in and Reliance on Algorithmic Advice," in *Proceedings of the 42nd International 2021*, Austin, Texas.

Schmitt, A., Zierau, N., Janson, A., and Leimeister, J. M. 2021b. "Voice as a contemporary frontier of interaction design," in *Proccedings of the 29th European Conference on Information Systems*, Virtual.

Schwartz, E. H. 2019. "64% of Americans Interested in Ordering Food Via Voice Assistant: Survey," available at https://voicebot.ai/2019/10/17/64-of-americans-interested-in-ordering-food-via-voice-assistant-survey/, accessed on Aug 24 2022.

Schwartz, E. H. 2022. "Amazon Brings Alexa In-Skill Purchasing to Canada," available at https://voicebot.ai/2022/04/13/amazon-brings-alexa-in-skill-purchasing-to-canada/, accessed on Apr 13 2022.

Schwarz, N., Jalbert, M., Noah, T., and Zhang, L. 2021. "Metacognitive experiences as information: Processing fluency in consumer judgment and decision making," *Consumer Psychology Review* (4:1), pp. 4-25 (doi: 10.1002/arcp.1067).

Sciuto, A., Saini, A., Forlizzi, J., and Hong, J. I. “"Hey Alexa, What's Up?": A Mixed-Methods Studies of In-Home Conversational Agent Usage,” in *Proceedings of the 2018 Designing Interactive Systems Conference*, pp. 857-868, Hong Kong, China (doi: 10.1145/3196709.3196772).

Shevechuk, N., Degimenci, K., and Oinas-Kukkonen, H. 2019. “Adoption of Gamified Persuasive Systems to Encourage Sustainable Behaviors: Interplay between perceived persuasiveness and cognitive absorption,” in *Proceedings of the 40th International Conference on Information Systems*, Munich, Germany.

Song, H., and Schwarz, N. 2009. “If it's difficult to pronounce, it must be risky,” *Psychological science* (20:2), pp. 135-138 (doi: 10.1111/j.1467-9280.2009.02267.x).

Sundar, S. S. 2008. “The MAIN model: A heuristic approach to understanding technology effects on credibility,” in *Digital media, youth, and credibility*, M. J. Metzger and A. J. Flanagin (eds.), Cambridge Mass.: MIT Press, pp. 73-100.

Tricks, H. 2021. “The return of one-to-one commerce,” available at https://www.economist.com/special-report/2021/03/11/the-return-of-one-to-one-commerce, accessed on May 1 2022.

Tsiros, M., and Heilman, C. M. 2005. “The Effect of Expiration Dates and Perceived Risk on Purchasing Behavior in Grocery Store Perishable Categories,” *Journal of Marketing* (69:2), pp. 114-129 (doi: 10.1509/jmkg.69.2.114.60762).

van den Oord, A., Kalchbrenner, N., Espeholt, L., Kavukcuoglu, K., Vinyals, O., and Graves, A. 2016. “Conditional Image Generation with PixelCNN Decoders,” *Advances in neural information processing systems* (29).

Viswanathan, M., and Childers, T. L. 1996. “Processing of Numerical and Verbal Product Information,” *Journal of Consumer Psychology* (5:4), pp. 359-385 (doi: 10.1207/s15327663jcp0504_03).

Wang, W., and Benbasat, I. 2007. “Recommendation Agents for Electronic Commerce: Effects of Explanation Facilities on Trusting Beliefs,” *Journal of Management Information Systems* (23:4), pp. 217-246 (doi: 10.2753/MIS0742-1222230410).

Whang, C., and Im, H. 2021. “"I Like Your Suggestion!" the role of humanlikeness and parasocial relationship on the website versus voice shopper's perception of recommendations,” *Psychology & Marketing* (38:4), pp. 581-595 (doi: 10.1002/mar.21437).

Winkielman, P., Schwarz, N., and Reber, R. 2003. “The Hedonic Marking of Processing Fluency: Implications for Evaluative Judgment,” in *The Psychology of Evaluation*, J. Musch and K. C. Klauer (eds.), Psychology Press, pp. 205-234 (doi: 10.4324/9781410606853-14).

Wirtz, J., Patterson, P. G., Kunz, W. H., Gruber, T., Lu, V. N., Paluch, S., and Martins, A. 2018. “Brave new world: service robots in the frontline,” *Journal of Service Management* (29:5), pp. 907-931 (doi: 10.1108/JOSM-04-2018-0119).

Xiao, and Benbasat. 2007. “E-Commerce Product Recommendation Agents: Use, Characteristics, and Impact,” *MIS Quarterly* (31:1), pp. 137-209 (doi: 10.2307/25148784).

Zellner, A. 1962. “An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias,” *Journal of the American Statistical Association* (57:298), pp. 348-368 (doi: 10.1080/01621459.1962.10480664).

Zierau, N., Hildebrand, C., Bergner, A., Busquet, F., Schmitt, A., and Marco Leimeister, J. 2022. “Voice bots on the frontline: Voice-based interfaces enhance flow-like consumer experiences & boost service outcomes,” *Journal of the Academy of Marketing Science* (doi: 10.1007/s11747-022-00868-5).

Zillmann, D., and Brosius, H.-B. 2012. *Exemplification in Communication*, Routledge.