The Role of Trusting Beliefs in Voice Assistants during Voice Shopping

Conference Paper · January 2021

2 authors, including:

Alex Mari
University of Zurich
9 PUBLICATIONS   **15** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project   Machine Learning in Marketing View project

The Role of Trusting Beliefs in Voice Assistants during Voice Shopping

Alex Mari
University of Zurich
alex.mari@uzh.ch

René Algesheimer
University of Zurich
rene.algesheimer@uzh.ch

Abstract

AI-based voice assistants such as Amazon Alexa deliver personalized product recommendations in order to match consumers' needs. The use of voice assistants for shopping purposes incorporates elements of risk affecting when and how they are considered trusted relationship partners. In this uncertain environment, 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), in the context of voice shopping. Findings of an individual-session online experiment (N=180) show a positive direct effect of trust on customer's satisfaction and a mediating role of set size, confirming consumers' bias towards default choices. This study highlights the consequences of trust in AI-enabled voice assistants for decision-making during utilitarian purchases.

1. Introduction

Over 200 million in-home voice assistants are installed globally [23], with a dominance of U.S. (Amazon Echo, Google Home) and China-based manufacturers (Alibaba Tmall Genie, Xiaomi Xiao AI). Voice assistants register the fastest development rate for a new technological medium, even above smartphones and tablets [8]. The term 'voice assistant' (VA) refers to conversational agents having the ability to self-improve their understanding of the interlocutor and context [33]. Using a combination of AI techniques, such as automatic speech recognition and natural language understanding, VAs can naturally converse with users, contextually elaborate requests, and dynamically expand their knowledge while learning from mistakes. As such, VAs develop a unique relationship with consumers bringing up a new set of interaction rules modeled after their active and proactive nature.

Besides simple commands such as playing music, providing weather information, and setting alarms, an increasing number of users are seeking more sophisticated experiences, such as shopping. In this respect, 21% of U.S. VA's owners have already made purchases using in-home VAs [23].

The act of placing orders on VAs using either native or third-party applications is referred to as 'voice commerce' or 'voice shopping'. These terms describe all the shopping activities allowing users to search for a product, listen to reviews, add items to a list, and track the order, etc. Shopping-related VAs can process orders with a simple command and without the user having to provide transactional information such as payment or address details. Consequently, they have the potential to substantially alter the process of product search and selection [6, 40, 32, 34]; however, little is known about consumers' decision-making in connection to the unique relationship users develop with VAs [42]. The social nature and uncommon choice framing of shopping-related VAs require new theories that have not yet been fully developed [8].

This paper examines how beliefs of trust towards an AI-enabled exchange partner, in the case of Amazon Alexa, affects satisfaction as a function of the number of alternatives sequentially recommended by the VA (consideration set size). Further, as VAs are designed such that only a single item is presented to a consumer at a time, this research explores the tendency of consumers to choose a default option. In the following paragraphs, we review background theories on product choice and framing. Next, we introduce a conceptual framework and describe the developed hypotheses. Finally, we discuss the experimental study findings before outlining opportunities for future research.

2. Theoretical background

A recommender agent generates personalized suggestions and attempts to match products to consumers' expressed preferences or implicit behaviors [50]. These algorithms are indispensable in online shopping environments where a potentially extensive set of alternatives are available. Research has shown that traditional recommender agents help consumers by

reducing consumer information overload and search complexity [20]. As a result, they have the potential to improve the quality of consumer decisions, which also increases consumer satisfaction and loyalty [11, 30].

In today's digital age, an increasing number of choices involve the use of AI-enabled agents. Powered by machine learning technologies, VAs are expected to match consumer preferences more closely than if they had chosen independently [1]. That is due to their ability to collect data systematically and silently over time [5]. As such, VAs can be conceptualized as interaction decision aid tools that promise fast, repeatable and low-cost decision-making combined with an increased level of accuracy, achieved through network effects and feedback loops [8, 20, 33].

2.1. Product choice during voice shopping

In the context of voice shopping, manufacturers like Amazon or Alibaba organize the general context in which people make decisions [56]. Amazon Alexa, the dominant product brokering VAs, is designed to find the best-suited products for consumers. Considering the worldwide market-leading position of Amazon as both a VA manufacturer and business-to-consumer e-commerce platform, our research focuses on the study of Alexa.

During product search on Alexa, the interaction flow with VAs changes whether the user wants to purchase in a product category for the first time, or repurchase a product in the same category [33]. In the first case, the dynamic dialog begins with an active decision by the user who needs to determine whether to search for a brand name (exact match), e.g. Pantene, or generic product category (broad match), e.g. shampoo. Alexa presents sequentially (versus simultaneously) a single option, a 'top search result', at a time. The assistant recommends new products only if the consumer answers "No" to the question, "Do you want to order this?". The purchasing process ends when a user agrees to purchase the item or quits the operation. Such a default does not force the user to make a decision; instead, it is more similar to a commonly used pre-checked box on internet forms. In the second case, when a user has already made a purchase using a VA, information stored in the system is retrieved to recommend a swift repurchase (automated match).

Our study focuses on a new product category purchase through broad search terms.

2.2. Choice framing and default bias

A consolidated body of research suggests that preferences are, on occasion, constructed in response to a decision [47]. When preferences are not articulated in the minds of decision-makers before starting a decision process, the trade-offs between options might not be clearly identified. As such, decisions can be profoundly affected by the context in which they take place.

The way a choice is presented to the decision-maker is described as 'choice framing' or 'choice architecture' [53, 55]. Voice assistants' manufacturers may influence a customer's environment so as to increase the likelihood that a particular option is chosen, even without changing incentives or prices [56]. For instance, they can influence choice by varying the presentation order of choice alternatives or attributes [27]. In times of product and information overload, individuals rely on simple heuristics to simplify their decision-making processes [60].

Increased difficulty of determining the best alternative raises the likelihood of choosing a pre-selected option or simply opting not to choose [51, 9]. One of the most powerful and consistently observed behavioral biases is the use of defaults. A default is the choice option that individuals adopt unless they actively choose an alternative [E.g., 2]. This human tendency to favor the status quo compared with other equally attractive decision alternatives is referred to as the 'default effect' [18]. Although defaults may dramatically impact consumer choice [E.g., 55], their unrecognized effects in some settings might turn them into 'hidden persuaders' [18]. The presence of a default option inflates its attractiveness, even when that default option is randomly assigned [27]. In that sense, defaults are often considered as nudges that influence consumers' decisions without restricting their freedom to choose [55].

A broader interpretation of the default effect includes the output of recommender agents [E.g., 21, 61], also in the context of voice commerce [34, 8].

3. Hypothesis development

Following the theory of reasoned action [13], we posit that shoppers' beliefs of trust towards VAs influence their choice and ultimately increases the consumer's satisfaction on the purchasing decision [14, 15, 22].
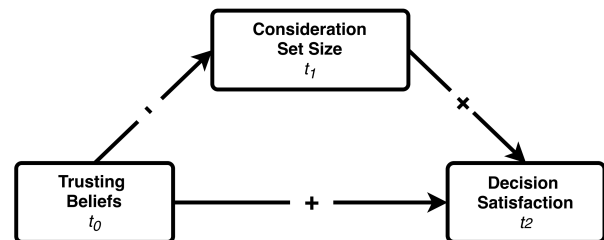

Figure 1. Conceptual framework.

Thus, the higher (lower) the trusting beliefs towards the VA and the smaller (larger) the consideration set size, i.e., the number of options the VA presents. Also, the smaller (larger) the option and the higher (lower) the decision satisfaction.

3.1. Main effect of trust on satisfaction

Trust is a defining attribute of the relationship between a VA and its users. Personal shoppers like Alexa or Genie exercise a greater influence when consumers are confident about their suggestions and in the process that generated them [19, 17]. Such a feeling of confidence towards an exchange partner incorporates elements of honesty, benevolence and competence. The "willingness to rely on an exchange partner in whom one has confidence" describes the multidimensional concept of trust [38, p.82]. Thus, trust in VA is one's belief that the other party will not behave opportunistically by taking advantage of the situation, instead it behaves in a fair and ethical manner, despite the trusting party's dependence and vulnerability. Given users' limited ability to comprehend the behavior of the shopping platform, often referred to as a 'black box' [49], all interactions with an intermediary actor like Alexa require an element of trust. Because of its nature, voice shopping implies an act of partial or total delegation to a VA, usually before the accuracy of product recommendations is actually experienced [29].

In this study, we embraced the initial conceptualization of trusting beliefs by Mayer, Davis, and Shoorman [35], utilized by McKnight, Choudhury, and Kacmar [37] in their work on e-commerce and then applied by Wang and Benbasat [58] to recommender agents and found to lead to trusting intentions [E.g., 36]. In line with McKnight et al. [37], trusting beliefs towards a VA include one's perceptions about the agent's competence, benevolence, and integrity. Competence is defined as the user's perception that a VA has the ability, skills and expertise to perform product recommendations effectively; benevolence belief is a user's perception that a VA cares about the consumer and acts in the consumer's interest; and, integrity belief is the perception that a VA adheres to a set of principles that users generally find acceptable (e.g., fair and unbiased recommendations).

Previous research established the relationship between trusting beliefs and satisfaction [39]. Although most studies concentrate on 'consumption satisfaction' [15], research shows that consumers experience satisfaction and dissatisfaction not only with the selected product but also with the purchase decision process itself. As such, satisfaction can be associated with the experience of a consumer arriving at a purchase decision, the so-called 'decision process satisfaction' [59].

Decision satisfaction refers to the level of feelings of satisfaction or regret about the chosen or rejected alternatives and in relation to the decision process [14, 15, 22]. Despite its relevance, the degree of satisfaction with the decision process has received little attention in marketing and information system research [E.g., 14]. Fitzsimons et al. [15] demonstrated the existence of decision and consumption satisfaction as conceptually distinct constructs. At the same time, they found a positive correlation between decision satisfaction and consumption satisfaction. Heitmann et al. [22] found that decision satisfaction not only contributes to consumption satisfaction but also has a positive effect on loyalty and willingness to recommend the product to others.

Collectively, these studies suggest that decision satisfaction, post-choice and pre-consumption represent a suitable construct to examine consumer response in experience-driven shopping environments like voice shopping. Thus, we advance that:

H1: Consumers' trusting beliefs (IV) towards a VA have a positive direct effect on decision satisfaction (DV).

3.2. Secondary effect of trust on default choice and consideration set size

Online defaults might be so well hidden that users may not be aware they had a choice in the first place. Defaults can exert significant effects not only in the presence of two alternatives but also as consumer choice expands. When Swedish citizens were asked to choose a retirement fund out of 456 alternatives, one-third of the participants invested solely in the default fund, despite having been encouraged to make active decisions [3]. Similar results were found with other consumer products such as personal computers, cars, treadmills, light bulbs and pizza [E.g., 52]. We posit that when consumers trust VAs, they are more inclined to opt for a default recommended option than those with lower beliefs of trust. Thus, we argue that:

H2a: Consumer's trusting beliefs towards a VA (IV) positively affect the tendency to select a default option.

Trust is the basis for task delegation. Recommender agents may replace traditional decision-making when consumers feel time constraints or recognize the referrer as a particularly knowledgeable source [45]. If consumers trust VAs to behave in their interests, they may be more likely to give up control in favor of convenience. We argue that, especially when the purchase involves utilitarian products such as

batteries, a belief of high trust (versus low trust) towards a VA corresponds to smaller (versus larger) options evaluated before making a purchase. In the presence of high (low) trust, consumers rely less (more) on brand and price signals in the decision. Thus, we posit that:

H₂ᵦ: Consumer's trusting beliefs (IV) negatively affect the consideration set size. This effect occurs independently of the brands (set brand) and prices (set price) offered.

3.3. Mediation effect of consideration set size

Prior research suggests that consideration set size can affect customer experience and resulting choice satisfaction [E.g., 11]. However, conflicting findings indicate that the effect of consideration set size varies across certain consumers and contexts. In the online environment, there is a negative effect of choice set size on satisfaction and choice confidence when users check facts on search engines from a shorter rather than longer list of results (6 versus 24) [46]. A similar effect is found in an online dating pool where users reported lower choice satisfaction when presented with a large set of potential partners versus a smaller pool [4]. Such a negative effect seems to be amplified under time constraints. Differently, other studies suggest that greater choice can increase the chance of preference matching and enhance perceptions of freedom of choice and control.

In the context of voice commerce, we argue that consideration set size functions as a mediator of the relationship between the beliefs of trust towards the VAs and the satisfaction about the shopping decision. Decision satisfaction was found to be largely influenced by the environment and choice architecture in which the decision occurs [61]. Thus, we posit that high (low) trust decreases (increases) the number of alternatives considered for shopping, which in turn, increases (decreases) the satisfaction about the shopping decision. In other words:

H₃: The number of options presented by the VA before the purchase (set size) mediates the relationship between trusting beliefs (IV) and decision satisfaction (DV). In particular, the greater the consideration set size and the lower is the decision satisfaction.

4. Experimental design

Product-brokering VAs like Alexa are designed such that only a single item is presented to a consumer at a time (default option). This paper explores the influence of choice framing on the trust users express towards the assistant. Drawing on recommender agents

and consumer's choice literature, an individual-session online experiment was conducted. Each subject made an actual voice purchase of a utilitarian product (batteries) using generic search terms (broad match). Batteries were chosen for their functional and low involvement nature. Further, several marketing experts consider this product category vulnerable to the rise of voice commerce [see 16, 54].

Purchasing data was collected using a third-party Alexa application 'Swiss Shopping', developed by the authors following systematic machine behavior observations [48, 33]. This Alexa 'skill' replicates the native voice shopping process in terms of flow, structure, and tone of the interaction, giving the user the feeling of dealing directly with Alexa throughout the entire task. As regular Alexa users are not yet able to distinguish Alexa's standard capabilities from third-party apps [31], a proprietary app that mimics Alexa's shopping functionalities represents a perfect opportunity to explore the choice framing effects in a controlled but realistic purchase environment.

The experiment described in the next paragraphs shows results in line with an exploratory pre-registered lab experiment with 30 subjects. The pilot test confirmed the correct manipulation of the treatment conditions and suggested minor adaptations to the study.

4.1. Participants

A total of 197 students were recruited through the university research service of two major universities in Zurich. Students fluent in English were invited to attend the study via email. Subjects were required to be in an undisturbed environment alone with a computer equipped with one camera and microphone and to have uninterrupted access to the Internet.

Students sample are often used in e-commerce research [57] and their demographic characteristics are in line with the main voice shopping user group [24]. Using software for block randomization (graphpad.com), participants were randomly assigned to one of three study conditions (high trust, low trust and no treatment).

A total of 180 subjects (60 per condition) were included in the analysis as they: i) showed-up to the video call on time, ii) passed the attention check at the beginning of the study (cognitive task), iii) made a purchase using Alexa without the researcher's support, iv) completed the study in more than 10 minutes.

Our respondents (*M*ₐgₑ=24) were nearly evenly distributed among the three experimental conditions in terms of gender (*Male*=44%), nationality (*Swiss*=42% vs. *Rest of world*, 38 nations) and experience with VAs. Sixty-six percent (n=119) of the study participants have never tried in-home VAs, while 12% (n=22) use it once

a month or more often. In terms of shopping behavior, a total of 8% (n=14) of respondents have made at least one purchase using VAs. Of those, only 3% (n=5) claimed to have purchased batteries using VAs in the past. Such percentages are in line with industry reports findings [see, 23]. The total study was on average completed in around 18.3 minutes (without briefing/debriefing), while the purchasing task lasted, on average, 2.2 minutes, with non-significant differences among conditions. Study participants received a standard compensation of 16.00 USD cash for a 30-minute commitment. In addition, they were offered to collect the purchased product.

4.2. Procedure and task

An experimental study applied a between-subjects design. A researcher welcomed individuals to the (Zoom) virtual room checking the fulfillment of all technical requirements. Five distinctive phases followed the initial greetings.

First, the study was introduced by saying: i) This is Alexa, a VA used for a variety of tasks such as checking the weather forecast, listening to music or setting up an alarm. For instance, (holding the device) I can just ask – Alexa, what's the weather like in Zurich? (Alexa's response is indented to demo a common feature of the device); ii) With Alexa you can also purchase household and grocery items, like shampoo or batteries. Today, we ask you to purchase a product using Alexa; iii) This study will last around 30 minutes. After reassuring the participants that he or she would be alone in the room with Alexa, the researcher shared a link on the Zoom chat containing an informed consent and study instructions. Finally, the researcher unmuted Alexa and left the room, closing the door behind.

Second, subjects were asked to read and reflect on a consumer report from a non-profit organization concerning the commerce capability of VAs like Alexa in terms of competence, benevolence and integrity [29, 37, 58]. Students read a report portraying either excellent or poor assistive abilities of VAs during shopping, depending on their assigned condition (high trust, low trust). The only differentiation between the documents submitted to the high and low trust groups consisted of the connotation of the sentences (positive versus negative). The control group (no treatment) was not exposed to any consumer report.

Third, the subjects were asked to purchase one packet of four AA batteries on Alexa for electronic devices such as a TV remote controls, clocks, or wireless mice. Individuals initiated the shopping capability on a 2nd generation Amazon Echo, the most popular VA device to date [23], saying, "Alexa, open Swiss Shopping", following which they entered a code

and asked for batteries. They were instructed to say "yes" when they wished to place an order and "no" if they wanted to hear more options.

Swiss Shopping includes 35 brands representing the 'top brands' of 'AA', 'alkaline', and 'single-use' batteries available on Amazon.com. As our study did not concern the decision-making of an optimal choice, so all the options were equally relevant, the order of the brand recommendation was randomized. Thus, all the brands, regardless of their popularity or market share, had the same chance to be recommended to the user. In order to eliminate the effect of quality and quantity, the items recommended by Alexa had the same product description (high-performance batteries) and quantity (pack of four). Also, in each session, every recommended brand was randomly associated with one of five price points representing the range between private label (cheapest) and national brand (premium) prices in both online and brick-and-mortar Swiss retailers (CHF 4.95, 5.45, 5.95, 6.45, 6.95). As such, brand name and price were the only elements changing among the available options.

Fourth, after the purchasing task was completed, respondents filled out a questionnaire on the same web page (Qualtrics). Finally, subjects were debriefed about the scope of the experiment and informed about the fictitious nature of the consumer report.

4.3. Measures

Study participants indicated on a questionnaire the extent to which they agreed with a total of 87 statements, grouped in 10 sessions, on a 7-point Likert scale (1=strongly disagree; 7=strongly agree).

Questions were related to the main examined constructs: trusting beliefs, only operationalized in the hierarchical regression analysis - paragraph 5.4 [adapted from 29, 37, 58] and decision satisfaction [14, 15, 22]. All core constructs and sub-constructs met the benchmark of good reliability ($\alpha > 0.80$) [43]. Secondary variables such as choice confidence, intention to adopt as a delegated agent or decision aid, post-purchase satisfaction decision, (future) intention to follow VA advice, were collected for exploratory purposes.

In addition, the survey assessed personality traits of the participants in terms of maximization, perceived satisficer and optimizer, propensity to default, propensity to trust technology, and propensity to trust objects / others. Along with the subjects' demographics, such as age, gender and nationality, we collected their product category knowledge – batteries, familiarity prior to the study with the product category, VAs, and voice commerce, and satisfaction with past experience (perceived usefulness). In addition, we investigated the

shopper's choice in terms of price and reason for default choice.

Additional measures: throughout the experiment, we collected extra measures from the log data of the Alexa skill. In particular, we stored the consideration set size, i.e., the number of recommendations presented by the VA before the user makes a purchase decision, consideration set brand, i.e., the percentage of brands known in the choice set prior to the experiment out of the total brands recommended by Alexa, consideration set price, i.e., the average price recommended by Alexa during the voice shopping session, and decision time. Extra information on pre-registration, manipulation and questionnaire can be found at https://cutt.ly/VAstudy.

5. Results

5.1. Manipulation check

A one-way analysis of variance (ANOVA) with multicategorical variables was conducted to compare trust beliefs scores in the high trust (M_{high}=4.66, SD=1.05), low trust (M_{low}=3.62, SD=1.05) and no treatment condition ($M_{notreat}$=4.13, SD=1.09). The difference in scores between the three groups was significantly different from zero at the 95% confidence interval, $F(2,177)$=14.339, p<.000. Participants in the high trust group have firmer trusting beliefs towards VAs than those in low trust group. Thus, we conclude that trust manipulation was successful.

Randomization check: a one-way ANOVA on the control variables excluded that the three conditions (coded: 0=no treatment, 1=high trust, 2=low trust) behave significantly different from each other. Thus, we refrain from discussing these variables further.

5.2. Main effect

A one-way ANOVA revealed that consumers' beliefs of trust affect how positively consumers are satisfied with their voice shopping decision, on average, $F(2,177)$=8.489, p<.000. In comparison to the no treatment condition ($M_{notreat}$=4.17, SD=1.14), shoppers in the high trust group show higher decision satisfaction (M_{high}=4.66, SD=1.18) than those in the low trust group (M_{low}=3.83, SD=1.10). Shoppers exposed to the high trust treatment are more satisfied with the shopping decision on average than those in the control group: b=0.514, $t(177)$=2.458, p=.015. Although only marginally significant (p=.104), those exposed to the low trust treatment are less satisfied with the shopping decision on average than those in the control group, b=-0.342, $t(177)$=-1.634. A 2-taled independent samples test reveals that high trust and low trust groups present

a decision satisfaction significantly different from each other, $t(118)$=4.096, p<.000. Thus, H_1 is supported.

5.3. Trust effect on the default option

There is a statistically significant association between conditions and default option $\chi^2(2)$=8.88, p=.012. The research hypothesis that differences in 'default option' are related to differences in 'trusting beliefs' is supported by this analysis. Such a relationship appears to have moderate strength (ϕc=.222).

As expected, shoppers in the high trust condition (M_{high}=0.60, SD=0.49) relied more on the recommendation provided by Alexa compared to the low trust group (M_{low}=0.47, SD=0.50). Surprisingly, the control group has made the most extensive use of default option (73%, n=60), compared to high trust (60%, n=60) and low trust (47%, n=60). A lack of attention in the no treatment group might be a sign of a "yeah, whatever" heuristic, especially popular in low involvement or utilitarian types of product choice [55].

Overall, 60% of respondents have purchased the first recommended brand and 83% (n=149) relied on the first three options provided by Alexa before finalizing the purchase (fig.2). The above analysis suggests that H_{2a} is supported.

5.4. Trust effect on consideration set size

There is a statistically significant association between conditions and consideration set size $\chi^2(2)$=14.830, p=.022. The research hypothesis that differences in 'consideration set size' is related to differences in 'trust' is supported by this analysis. Such a relationship appears to have moderate strength (ϕc=.205). Thus, H_{2b} is supported.
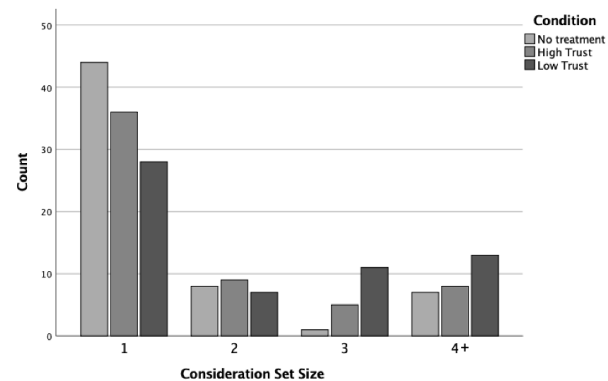

Figure 2. Crosstabulation: Set size * Condition.

Furthermore, to explore the effect of trusting beliefs (12-item construct) on the n. of options recommended and the relative effect of the VA's

recommended brands and price, we conducted a hierarchical multiple regression analysis. Prior to the statistical analysis that evaluates the prediction of consideration set size (DV) from consideration set price, consideration set brand and trusting beliefs (IVs), the relevant assumptions were tested. The assumption of singularity and multicollinearity were met [12]. Extreme univariate outliers identified in initial data screening were modified. Residual and scatter plots indicated the assumptions of normality, linearity and homoscedasticity were all satisfied [12].

DV: Consideration set size	β	t	sr²	R	R²	ΔR²
Step 1				.00	.00	.00
- Consideration set price	.00	-0.00	.00			
Step 2				.18	.03	.03
- Consideration set price	*.00*	*.10*	*.00*			
- Consideration set brand	-0.19	-2.50*	-0.19			
Step 3				.39	.15	.12
- Consideration set price	*.01*	*.17*	*.01*			
- Consideration set brand	*-0.16*	*-2.27**	*-0.17*			
- Condition: High trust	.14	1.67	.13			
- Condition: Low trust	.19	2.26*	.17			
- Trusting beliefs (12-item construct)	-0.28	-3.61***	-0.27			

*N=177; *p <.05, **p <.01, ***p<.001*

Table 1. Summary of Hierarchical Regression Analysis.

In a three-stage hierarchical multiple regression, the variables were entered to disentangle the effect of consideration set price, consideration set brand, from trusting belief, the core variable in our study (tab.1). During stage one, it emerged that consideration set price does not contribute significantly to the regression model, $F(1,175)$=.00, p=.997. Introducing the consideration set brand variable explained 3.5% of the variation in consideration set size and this change in R^2 was significant, $F(1,174)$=6.27, p=.013. Adding trusting beliefs and accounting for conditions, our regression model explained an additional 11.4% of the variation in consideration set size and this change in R^2 was significant, $F(3, 171)$=7.60, p<.000.

Together, the three independent variables and the condition variables (high trust, low trust, control) accounted for 14.8% of the variance in set size. The regression coefficients show that both trusting beliefs [β=-.278, 95% CI (-0.637, -0.187), p<.000] and consideration set brand [β=-.162, 95% CI (-0.012, -0.001), p=.025] contribute to negatively affect consideration set size. In other words, the higher (lower) the trust towards Alexa as well as the higher (lower) the percentage of known brands in the recommended set, the lower the number of options the VA is asked to recommend before the purchase. However, consideration set price remained a non-significant predictor of consideration set size throughout the three stages of the regression model. Thus, our assumption of

trusting belief effect on set size independent of brands and prices, was confirmed for set price and rejected for set brand.

5.5. Simple mediation

A simple mediation analysis was used to estimate and test hypotheses about the causal influence paths from trusting beliefs to decision satisfaction, the first through the proposed mediator, consideration set size, and the second independent of the X→M→Y mechanism (PROCESS: Model 4; Hayes, 21). We treated the independent variable as multicategorical using the option: mcx. Multiple regression analysis was conducted to assess each component of the hypothesized mediation model (fig.3).
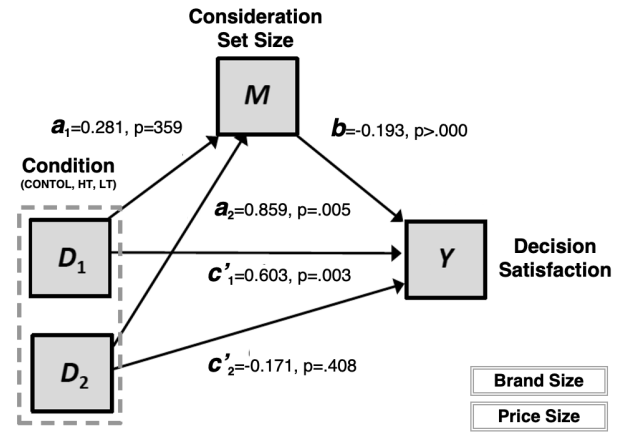

Figure 3. Statistical model.

A one-way ANOVA revealed that consumers' beliefs of trust affect the average number of options the VA presents, $F(2,174)$=4.120, p=.018. Relative to the control group ($M_{notreat}$=1.65, SD=1.42), the low trust group evaluated significantly more product recommendations from Alexa, on average (b=0.859, p=.005). Although marginally significant, $t(115)$=-1.764, p=.080), shoppers in the low trust group show higher consideration set (M_{low}=2.51, SD=1.90) than those in the high trust group (M_{high}=1.93, SD=1.63).

When considering the mediating effect of consideration set size, compared to the control group, shoppers in the treated groups are on average less satisfied about their shopping decisions. In particular, controlling for consideration set size, shoppers in the high trust group show higher decision satisfaction on average than those in the control condition (b=0.603, p=.003). Also, holding condition constant, those who have evaluated more options before making a shopping

decision show lower decision satisfaction (b=-0.193, p<.000).

Compared to the control group, when a shopper does not trust Alexa he or she is 0.165 units less satisfied as a result of an increase in consideration set size. In other words, a lower (versus higher) level of trust in Alexa brings shoppers to review more options, which in turn reduces decision satisfaction. The relative indirect effects of low trust compare to no treatment is significantly different from zero (relative indirect effect=-0.166, 95% *CI:* -0.325 to -0.450). There is no direct effect of low trust on decision satisfaction.

The relative indirect effects of high trust relative to no treatment is not significantly different from zero (relative indirect effect=-0.054, 95% *CI:* -0.176 to 0.049). Given that at least one relative indirect effect is different from zero, we have evidence that the effect of trusting beliefs on decision satisfaction is operating, at least in part, through consideration set size [21]. Furthermore, a test of the omnibus direct effect supports that the three conditions differ on average in how positively shoppers are satisfied after accounting for group differences in consideration set size, R^2change=0.077, F(2,177)=7.961, p<.000.

The results described above hold true after controlling for the effects of brand (knowledge of purchased brand - y/n; consideration set brand) and price (price paid; consideration set price). Thus, we conclude that H_3 is supported.

6. Discussion and conclusion

This research explores the effect of trust on decision satisfaction through the mediating effect of consideration set size. In particular, we experimentally study how Alexa's choice framing affects consumers' satisfaction during voice shopping. To the authors' knowledge, this is the first study investigating consumer choices through an actual voice commerce purchase while using an ad-hoc designed voice skill. This study contributes in several ways to the understanding of human relationships with AI-based agents and their effect on decision making.

First, results confirm that consumers' beliefs of trust in a VA influence their satisfaction with the shopping decision (H_1). Purchase decisions involving a new shopping method with unique characteristics are inherently risky. The development of consumer trust is a critical driver of decision satisfaction. As such, the process through which a decision is made, rather than its outcome, constitutes an important aspect of satisfaction in the voice shopping environment.

Second, our analysis underlines the strategic importance of defaults in product choice involving VAs while bringing evidence that trust negatively affects the

n. of options a VA presents (H_2). Although choice framing does not force the user to make a decision, this may produce several unanticipated effects. A simplified representation of the marketplace reduces consumer visibility of alternative products and features ('echo-chamber' effect), while increasing brand polarisation [34]. On the positive side, a lower number of alternatives presented may lead to less regret over foregone options.

Third, we bring evidence that the effect of trusting beliefs on decision satisfaction is operating, at least in part, through consideration set size (H_3). Voice assistants are set to increasingly influence consumer behavior as they become better at proactively learning users' preferences and habits. Where a lower n. of options recommended by VAs corresponds to higher decision satisfaction, an increase in consumer trust has the potential, over time, to significantly reduce (or even eliminate) the need for human decision making [7]. Consequently, the mediating effect of set size plays an important role in the future diffusion of voice commerce.

Finally, differently from other studies [41], we report limited importance of brand (compared to trust) and the non-existing effect of price on consideration set size. Especially when individuals do not possess brand preferences in a determined category, we can expect a greater influence of Alexa's choice framing. In fact, when users shop on Alexa using generic terms like toilet paper or razor, they may construct preferences on-the-go, relying on the VA recommendations more.

Several limitations of the proposed study should be recognized. First, this research was conducted within the context of the initial trust timeframe. This static perspective fails to consider that experienced customers may exhibit different trusting behaviors as the factors influencing trust might become more or less salient over time. Thus, future empirical investigations should employ a longitudinal design to explore the process by which trust evolves. Second, our study is limited to the trust of a specific trustor (human) for a specific trustee (VA). This dyadic perspective does not consider the role of the VA's manufacturer or voice shopping provider, i.e., Amazon. Future studies should consider consumers' perceptions towards all key actors in the shopping ecosystem. Third, the purchase of only one type of product limits the generalizability of the findings. Further, individuals in our study did not express the intent to purchase batteries before the experiment. Future studies need to explore differential dynamics between product categories (e.g., utilitarian versus hedonic) after purchase intent in those categories is expressed. Finally, upcoming research may analyze the interplay of trusting beliefs' sub-constructs (competence, benevolence, integrity) and introduce the dimension of emotional trust next to cognitive [28].

7. References

[1] André, Q., Z. Carmon, K. Wertenbroch, A. Crum, D. Frank, W. Goldstein, ... and H. Yang, "Consumer choice and autonomy in the age of artificial intelligence and big data", Customer Needs and Solutions, 5(1-2), 2018, pp. 28–37.

[2] Brown, C.L., and A. Krishna, "The skeptical shopper: A metacognitive account for the effects of default options on choice", Journal of Consumer Research, 31(3), 2004, pp. 529–539.

[3] Cronqvist, H., and R.H. Thaler, "Design choices in privatized social-security systems: Learning from the Swedish experience", American Economic Review, 94(2), 2004, pp. 424–428.

[4] D'Angelo, J.D., and C.L. Toma, "There are plenty of fish in the sea: The effects of choice overload and reversibility on online daters' satisfaction with selected partners", Media Psychology, 20(1), 2017, pp. 1–27.

[5] Davenport, T., A. Guha, D. Grewal, and T. Bressgott, "How artificial intelligence will change the future of marketing", Journal of the Academy of Marketing Science, 48(1), 2020, pp. 24–42.

[6] Dawar, N., and N. Bendle, "Marketing in the age of Alexa", Harvard Business Review, 96(3), 2018, pp. 80–86.

[7] De Bellis, E., and G.V. Johar, "Autonomous Shopping Systems: Identifying and Overcoming Barriers to Consumer Adoption", Journal of Retailing, 96(1), 2020, pp.74–87.

[8] Dellaert, B.G., S.B. Shu, T.A. Arentze, T. Baker, K. Diehl, B. Donkers, … and H. Oppewal, "Consumer Decisions With Artificially Intelligent Voice Assistants", Marketing Letters, 2020, pp.1–13.

[9] Dhar, R., "Consumer preference for a no-choice option", Journal of Consumer Research, 24(2), 1997, pp. 215–231.

[10] Diehl, K., and C. Poynor, "Great expectations?! Assortment size, expectations, and satisfaction", Journal of Marketing Research, 47(2), 2010, pp. 312–322.

[11] Diehl, K., L.J. Kornish, and J.G. Lynch Jr, "Smart agents: When lower search costs for quality information increase price sensitivity", Journal of Consumer Research, 30(1), 2003, pp. 56–71.

[12] Field, A., Discovering statistics using SPSS: (and sex and drugs and rock'n'roll), Sage, 2009.

[13] Fishbein, M., and I. Ajzen, Belief, attitude, intention, and behavior: An introduction to theory and research, 1977.

[14] Fitzsimons, G.J. "Consumer response to stockouts", Journal of Consumer Research, 27(2), 2000, pp. 249–266.

[15] Fitzsimons, G.J., E.A. Greenleaf, and D.R. Lehmann, "Decision and consumption satisfaction: Implications for channel relations", Marketing Studies Center Working Paper Series, 313. 1997.

[16] Galloway, S., "This Technology Kills Brands" – www.youtube.com/watch?v=BXEu8RcneZQ, accessed 04-08-2020.

[17] Gefen, D., E. Karahanna, and D.W. Straub, "Trust and TAM in online shopping: An integrated model", MIS Quarterly, 27(1), 2003, pp. 51–90.

[18] Goldstein, D.G., E.J. Johnson, A. Herrmann, and M. Heitmann, "Nudge your customers toward better choices", Harvard Business Review, 86(12), 2008, pp. 99–105.

[19] Häubl, G., and K.B. Murray, "Preference construction and persistence in digital marketplaces: The role of electronic recommendation agents", Journal of Consumer Psychology, 13(1-2), 2003, pp. 75–91.

[20] Häubl, G., and V. Trifts, "Consumer decision making in online shopping environments: The effects of interactive decision aids", Marketing Science, 19(1), 2000, pp. 4–21.

[21] Hayes, A.F., Introduction to mediation, moderation, and conditional process analysis: A regression-based approach, Guilford publications, 2017.

[22] Heitmann, M., D.R. Lehmann, and A. Herrmann, "Choice goal attainment and decision and consumption satisfaction", Journal of Marketing Research, 44(2), 2007, pp. 234–250.

[23] https://www.emarketer.com/chart/230338, accessed: 18-06-2019.

[24] https://voicebot.ai/2018/07/12/, accessed: 15-09-2020.

[25] https://voicebot.ai/2019/04/15, accessed: 15-04-2019.

[26] Jannach, D., M. Zanker, A. Felfernig, and G. Friedrich, Recommender systems: an introduction, Cambridge University Press, 2010.

[27] Johnson, E.J., J. Hershey, J. Meszaros, and H. Kunreuther, "Framing, probability distortions, and insurance decisions", Journal of Risk and Uncertainty, 7(1), 1993, pp. 35–51.

[28] Komiak, S.X., and I. Benbasat, "Understanding customer trust in agent-mediated electronic commerce, web-mediated electronic commerce, and traditional commerce", Information Technology and Management, 5(1-2), 2004, pp.181–207.

[29] Komiak, S.X., and I. Benbasat, "The effects of personalization and familiarity on trust and adoption of recommendation agents", MIS Quarterly, 2006, pp. 941–960.

[30] Liang, T.P., H.Y. Chen, and E. Turban, "Effect of personalization on the perceived usefulness of online customer services: A dual-core theory", in Proceedings of the 11th International Conference on Electronic Commerce, pp. 279–288. 2009.

[31] Major, D.J., D.Y. Huang, M. Chetty, and N. Feamster, "Alexa, Who Am I Speaking To? Understanding Users' Ability to Identify Third-Party Apps on Amazon Alexa", arXiv preprint arXiv:1910.14112. 2019.

[32] Mari, A. "Voice Commerce: Understanding shopping-related voice assistants and their effect on brands", in IMMAA Annual Conference. Northwestern University in Qatar, Doha (Qatar). 2019.

[33] Mari, A., A. Mandelli, and R. Algesheimer, "The evolution of marketing in the context of voice commerce: A managerial perspective", in International Conference on Human-Computer Interaction, Springer, Cham, pp. 405–425. 2020.

[34] Mari, A., and R. Algesheimer, "Unfolding AI-based Voice Assistants for Digital Marketing, Preparing for

Voice Marketing and Commerce", In Contemporary Issues in Digital Marketing, Taylor & Francis. 2021.

[35] Mayer, R.C., J.H. Davis, and F.D. Schoorman, "An integrative model of organizational trust", Academy of Management Review, 20(3), 1995, pp. 709–734.

[36] McKnight, D.H., and N.L. Chervany, "What trust means in e-commerce customer relationships: An interdisciplinary conceptual typology", International Journal of Electronic Commerce, 6(2), 2001, pp. 35–59.

[37] McKnight, D.H., V. Choudhury, and C. Kacmar, "Developing and validating trust measures for e-commerce: An integrative typology", Information Systems Research, 13(3), 2002, pp. 334–359.

[38] Moorman, C., R. Deshpande, and G. Zaltman, "Factors affecting trust in market research relationships", Journal of Marketing, 57(1), 1993, pp. 81–101.

[39] Morgan, R.M., and S.D. Hunt, "The commitment-trust theory of relationship marketing", Journal of Marketing, 58(3), 1994, pp. 20–38.

[40] Munz, K., and V. Morwitz, "Not-so Easy Listening: Roots and Repercussions of Auditory Choice Difficulty in Voice Commerce", Available at SSRN 3462714. 2019.

[41] Nordheim, C.B., A. Følstad, and C.A. Bjørkli, "An initial model of trust in chatbots for customer service—Findings from a questionnaire study", Interacting with Computers, 31(3), 2019, pp.317–335.

[42] Novak, T.P., and D.L. Hoffman, "Relationship journeys in the internet of things: a new framework for understanding interactions between consumers and smart objects", Journal of the Academy of Marketing Science, 47(2), 2019, pp. 216–237.

[43] Nunnally, J.C., Psychometric theory 3E. Tata McGraw-hill education, 1994.

[44] Oliver, R.L., "A cognitive model of the antecedents and consequences of satisfaction decisions", Journal of Marketing Research, 17(4), 1980, pp. 460–469.

[45] Olshavsky, R.W., and D.H. Granbois, "Consumer decision making—fact or fiction?", Journal of Consumer Research, 6(2), 1979, pp. 93–100.

[46] Oulasvirta, A., J.P. Hukkinen, and B. Schwartz, "When more is less: the paradox of choice in search engine use", In Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval (pp. 516–523). 2009.

[47] Payne, J.W., J.R. Bettman, and E.J. Johnson, "Behavioral decision research: A constructive processing perspective", Annual Review of Psychology, 43(1), 1992, pp. 87–131.

[48] Rahwan, I., M. Cebrian, N. Obradovich, J. Bongard, J.F. Bonnefon, C. Breazeal, ... and N.R. Jennings, "Machine behaviour", Nature, 568(7753), 2019, pp. 477–486.

[49] Rai, A., "Explainable AI: from black box to glass box", Journal of the Academy of Marketing Science, 48(1), 2020, pp. 137–141.

[50] Resnick, P., and H.R. Varian, "Recommender systems", Communications of the ACM, 40(3), 1997, pp. 56–58.

[51] Samuelson, W., and R. Zeckhauser, "Status quo bias in decision making", Journal of Risk and Uncertainty, 1(1), 1988, pp. 7–59.

[52] Shafir, E., I. Simonson, and A. Tversky, "Reason-based 6choice", Cognition, 49(1-2), 1993, pp. 11–36.

[53] Smith, S.M., and I.P. Levin, "Need for cognition and choice framing effects", Journal of Behavioral Decision Making, 9(4), 1996, pp. 283–290.

[54] Sterne, J., Artificial intelligence for marketing: practical applications, John Wiley & Sons, 2017.

[55] Thaler, R.H., and C.R. Sunstein, Nudge: Improving decisions about health, wealth, and happiness, Penguin, 2009.

[56] Thaler, R.H., C.R. Sunstein, and J.P. Balz, "Choice architecture", The Behavioral Foundations of Public Policy, 2013, pp. 428–439.

[57] Walczuch, R., and H. Lundgren, "Psychological antecedents of institution-based consumer trust in e-retailing", Information & Management, 42(1), 2004, pp.159–177.

[58] Wang, W., and I. Benbasat, "Recommendation agents for electronic commerce: Effects of explanation facilities on trusting beliefs", Journal of Management Information Systems, 23(4), 2007, pp. 217–246.

[59] Westbrook, R.A., J.W. Newman, and J.R. Taylor, "Satisfaction/Dissatisfaction in the Purchase Decision Process: Are consumers really as dissatisfied as many studies claim?", Journal of Marketing, 42(4), 1978, pp. 54–60.

[60] Wright, P., "Consumer choice strategies: Simplifying vs. optimizing", Journal of Marketing Research, 12(1), 1975, pp. 60–67.

[61] Xiao, B., and I. Benbasat, "E-commerce product recommendation agents: use, characteristics, and impact", MIS Quarterly, 31(1), 2007, pp. 137–209.