Rising like a Phoenix: Emerging from the Pandemic and Reshaping Human Endeavors with Digital Technologies ICIS 2023

Human Technology Interaction

Dec 11th, 12:00 AM

AI-based Technologies for Everyone: How and Why to Adapt Voice Assistants’ Complexity to Older Adults

Christine Rzepka
Ludwig-Maximilians-Universität München, rzepka@bwl.lmu.de

Benedikt Berger
University of Münster, benedikt.berger@ercis.uni-muenster.de

Stefan Tams
HEC Montreal, stefan.tams@hec.ca

Recommended Citation

Rzepka, Christine; Berger, Benedikt; and Tams, Stefan, "AI-based Technologies for Everyone: How and Why to Adapt Voice Assistants’ Complexity to Older Adults" (2023). *Rising like a Phoenix: Emerging from the Pandemic and Reshaping Human Endeavors with Digital Technologies ICIS 2023*. 16.
https://aisel.aisnet.org/icis2023/hti/hti/16

AI-based Technologies for Everyone: How and Why to Adapt Voice Assistants' Complexity to Older Adults

Completed Research Paper

Christine Rzepka
LMU Munich
Ludwigstr. 28, 80539 Munich, Germany
ch.rzepka@lmu.de

Benedikt Berger
University of Münster
Leonardo-Campus 11, 48149 Münster, Germany
benedikt.berger@ercis.uni-muenster.de

Stefan Tams
HEC Montréal
3000 Chemin de la Côte-Sainte-Catherine, Montréal, Canada H3T 2A7
stefan.tams@hec.ca

Abstract

Technological advancements in the area of artificial intelligence have rapidly improved the performance of speech recognition and natural language processing. These improvements have facilitated the proliferation of voice assistants (VAs), which can understand human speech and provide spoken answers to assist in various tasks. More and more individuals and organizations adopt VAs because they value the naturalness of speech interaction. However, speech interaction is of ephemeral nature and processed in sequential order, which puts cognitive load on the user. Therefore, we investigate the relationship between the complexity of speech interaction and the interaction outcomes enjoyment, satisfaction, and intention to explore. Our results show that this relationship has an inverted U-shape for people with above-median information processing speed (i.e., younger adults) but is negatively linear otherwise. The results contribute to the literature on interface complexity and on the use of IT systems by the elder.

Keywords: Voice assistants, speech interaction, age, information processing speed

Introduction

Technological advancements in artificial intelligence (AI) have rapidly improved the performance of speech recognition and natural language processing (Hirschberg and Manning 2015). As a consequence, more and more households and organizations adopt voice assistants (VAs), which understand human speech and provide spoken answers to assist in various tasks (Hoy 2018). As the technology continues to improve, the tasks VAs can solve and the required interactions between VAs and users become more complex. Following the release of the conversational text-generator ChatGPT, for instance, first companies experiment with integrating ChatGPT in VAs to respond to multiple questions and prompts at once (Tuohy 2023). Taking advantage of VAs requires a thorough understanding of how humans interact with them. Therefore, the applicability of speech interaction to particular tasks, users, and contexts is an important topic for information systems (IS) research (Diederich et al. 2022).

While speech interactions with information technology (IT) systems are often perceived as being more natural than interactions via other modalities (Kock 2004; Schmitt et al. 2021), more complex speech interactions could pose a problem: information presented in a spoken way is of ephemeral nature and processed in sequential order, which puts cognitive load on the user (Le Bigot et al. 2004; Rubin et al. 2000). Therefore, users need to memorize the given information, which gets more difficult the more information is exchanged in the conversation (Miller 1994). These findings stress the important role of complexity in the design of conversations with VAs.

Extant research on the complexity of interfaces shows that finding the right balance of complexity is a determinant for user's assessment of the interaction with IT systems (Ghasemaghaei et al. 2019; Nadkarni and Gupta 2007). Studies on website design find that the number of links, pictures, and the website's length significantly raise the cognitive load when interacting with the website and, thus, the website's perceived complexity (Geissler et al. 2001; Nadkarni and Gupta 2007; Wang et al. 2014). These studies also find that a certain degree of complexity can be beneficial. For example, product comparisons involving only few product attributes are less complex but also provide less information, making the purchasing process less accurate. Complexity can, thus, have positive effects as it comes with higher information richness (Ghasemaghaei et al. 2019). Accordingly, multiple studies find an inverted U-shaped relationship with an optimum ('sweet spot') in user satisfaction at a moderate level of complexity (Geissler et al. 2001; Nadkarni and Gupta 2007). These studies, however, examine visual information presentation and are, therefore, not applicable to purely spoken interactions. To prevent users from terminating conversations with VAs, this study seeks to understand the determinants and effects of complexity for speech interaction with VAs. Therefore, we pose the following research question:

RQ1: How and why does VAs' interaction complexity affect interaction outcomes?

Additionally, it stands to reason that the interaction complexity of VA use is not only a technical matter but also depends on users (Bardhan et al. 2020; Diederich et al. 2022). Extant research suggests that an IT system's complexity is among the strongest inhibitors of use by older adults (Galliers et al. 2012). Because cognitive resources, like information processing speed, decline with age, excessive complexity is particularly problematic for older users as their cognitive capacity is limited (Salthouse 1996; Tams 2022). Building on those insights, we suppose that the 'sweet spot' between interaction complexity and interaction outcomes shifts towards less complex interfaces for the elder generation. This raises the question whether and how adapting speech interactions to older users facilitates their use of VAs. Investigating the effect of age on the interaction with VAs is important because VAs' hands- and eyes-free interaction capabilities are particularly beneficial for older users, who may suffer from physical impairments. Additionally, adults over the age of 65 are expected to make up 28% of the worldwide population in 2050 (OECD 2017). However, system designers are often younger and design applications according to their own needs and experiences (Hawthorn 2007), potentially neglecting characteristics of other user groups (Fisk et al. 2020). At the same time, new technological advancements offer the opportunity to identify VA users' age and to adapt the interaction accordingly (Burkhardt et al. 2021; Tóth et al. 2018). Therefore, we add the question:

RQ2: Does a user's age impact the relationship between VAs' interaction complexity and interaction outcomes?

To address these research questions, we developed a research model based on stimulus complexity theory (Berlyne 1960), the theory of processing speed (Salthouse 1996), and the theory of compensatory adaptation (Kock 1998; Kock 2001). We tested the model in a between-subject online experiment. In this experiment, participants planned a trip using a VA, which elicited twelve information cues for successful task completion (e.g., destination, departure date). Using a factorial design, we compared four conditions that differed in the VA's interaction complexity, which we operationalized as the number of information cues per interaction turn. The results from our main experiment with 405 participants show that users' age-related deficits in their information processing speed changes the nature of the relationship between perceived interaction complexity and the interaction outcomes satisfaction, enjoyment, and intention to explore. We contribute to research on adaptable AI-enabled systems as well as on the use of IT systems by the elder, showing that even supposedly simple and natural ways of interaction need to account for the specific characteristics of older adults. Our findings further inform practitioners, who should adapt the interaction complexity of their VA's conversational design to users' age.

Theoretical Background

Stimulus Complexity Theory and Conversational Design

We draw on stimulus complexity theory (Berlyne 1960) as foundation for our hypotheses because this theory explains the effects of IT systems' complexity on information processing. In general, complexity describes "the amount of variety or diversity in a stimulus pattern" (Berlyne 1960, p. 38) and increases with the number of distinguishable elements and the dissimilarity between elements. Berlyne's theory of stimulus complexity assumes an inverse U-shaped relationship between an IT system's complexity and communication effectiveness (i.e., the interaction outcomes desired from the interaction). If complexity is too low, users are bored from using the IT system and miss out on information richness (Hall and Hanna 2004; Palmer 2002). If complexity is too high, users suffer from cognitive overload, leading to frustration (Venkatesh et al. 2003). Hence, medium complexity levels should be desirable, which has also been confirmed for the design of websites, leading to improved satisfaction (Geissler et al. 2001; Nadkarni and Gupta 2007). Research on website complexity defines complexity as the amount of information presented on a website, e.g., in form of pictures, text, or links. Compared to the permanent nature of text-based interaction, however, speech interactions are characterized by ephemeral information, which is presented sequentially (Rubin et al. 2000; Schmitt et al. 2021). This distinct way of interacting with IT systems thus requires a new conceptualization of complexity for speech interactions.

In linguistics, interaction is defined as the "interplay between what one speaker is doing in a turn-at-talk and what the other did in their prior turn" (Drew 2013, p. 131). Thus, a conversation consists of several turns, which describe sequential units of conversational communication. Turns switch sequentially from one speaker to the other, a process which is called "turn-taking" (Wilson et al. 1984, p. 160). Each speaker grounds his or her response in the previous turn to ensure conversational contingency (Clark 1996). Therefore, each speaker starts planning his response for the following turn already while listening to the current speaker. This process becomes more complex and requires more cognitive effort the more information is given and requires a proper response. Compared to natural interactions, which do not have a prescribed order, length, or specific content of particular turns (Wilson et al. 1984), VAs converse in a more structured manner. They ask for predefined parameters to solve the task for which they are built and thus control the interaction (Drew 2013; McTear 2018). In this context, a higher number of turns can be positively related to users' interaction experience, whereas a smaller number allows for a fast and efficient, but more complex conversation (Jain et al. 2018). However, users' preferences in light of this trade-off remain mostly unexplored (Diederich et al. 2022). Such knowledge becomes especially important when considering that VAs evolve from being able to process one-shot queries to leading open-ended dialogues with the user (McTear 2018; Vakulenko et al. 2020).

Previous studies on complexity in the context of conversational agents (including chatbots and VAs) mostly investigate the complexity of the task to be solved and find that conversational agents suit tasks of lower complexity more than tasks of higher complexity (Hsu et al. 2023; Kiseleva et al. 2016). Schick et al. (2022) find chatbot use itself to be more complex than the use of a paper-and-pencil or a web-based form for mental health assessment. However, we are not aware of any study that investigates different degrees of interaction complexity of conversational agents, let alone VAs. Since the way in which a VA guides the conversation can make the interaction seem more or less complex for the user, we aim to understand the impact of conversational structure in terms of turn-taking design on interaction outcomes.

The Impact of Users' Age and Compensatory Adaptation

How complex a user perceives the interaction with a VA to be also depends on the user's characteristics. Because VAs can adapt their conversational design to their users (Diederich et al. 2022), it is worthwhile to investigate how individual user characteristics affect the consequences of VAs' interaction complexity. Specifically for older adults, VAs offer tremendous benefits because of the physical and cognitive effects of ageing that influence how older adults use computers (Bardhan et al. 2020; Kowalski et al. 2019). Since VAs enable their users to interact hands- and eyes-free in a natural way, older adults are not constrained by physical impairments of eyesight or motor skills. However, spoken interaction itself puts cognitive demand on the user, i.e. for formulating a proper command, for comprehending synthesized voices, and for processing spoken information (Craik and Byrd 1982). The theory of processing speed (Salthouse 1996) explains that this effect is due to reductions in brain volume and in cerebral blood flow, which slows down neural impulses (Park and Festini 2017). Therefore, conversations with VAs may be cognitively more difficult for older than for younger adults and require adaptations in terms of conversational structure and flow to ease their use of new technologies (Bardhan et al. 2020).

Initial research on elderly's use of VAs confirms that age-related cognitive decline is associated with certain vocal characteristics and behavioral differences compared to younger users, such as pauses and hesitations during conversations with VAs (Oh et al. 2020). For solving complex tasks, such as seeking health information, older adults expect VAs to proceed in a step-by-step manner (Gollasch and Weber 2021)and prefer to use other means than VAs (Brewer et al. 2022) if possible. Islam and Chaudhry (2023) find that complex questions and cognitive overload are common concerns among older users of VAs. However, receiving training in the use of VAs improves older adults' long-term VA use and feature exploration (Upadhyay et al. 2023).

Besides training, VAs may compensate for older adults' (decreasing) cognitive abilities by adapting the interaction complexity dynamically (Bardhan et al. 2020; Diederich et al. 2022). Serrano and Karahanna (2016) draw on the theory of compensatory adaptation (Kock 1998; Kock 2001) to examine how task-specific user capabilities can overcome limitations of IT systems. According to this theory, using IT systems can have positive effects on interaction outcomes even if the IT system only offers imperfect communication features. In contrast, our study aims to understand how IT systems can overcome limited user abilities, even if this requires limiting the IT system's communicative features. More specifically, we propose that adapting the level of an IT system's complexity may compensate for users' age-related deficits. Therefore, we assume that older adults experience the same inverted U-shaped relationship between an IT system's complexity and interaction outcomes as originally proposed by Berlyne (1960). However, this curve is shifted towards lower complexity levels (see Figure 1), compared to the younger population. Our study seeks to investigate this relationship between VA's interaction complexity, interaction outcomes, and users' age.

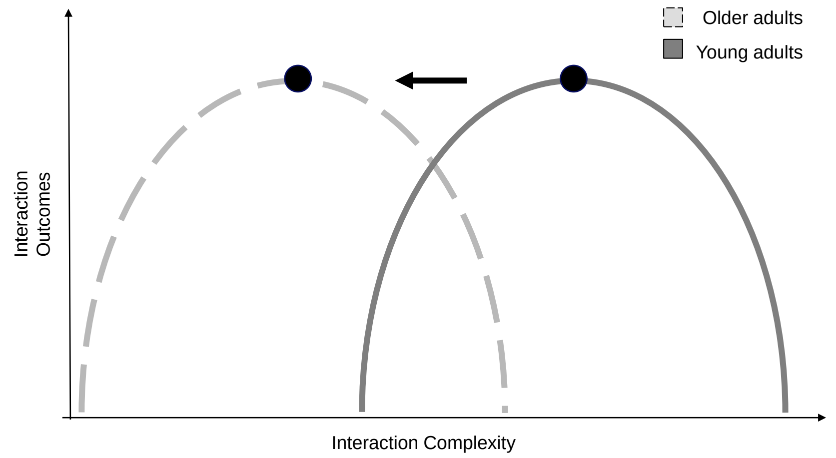


Figure 1. Conceptual Model

Research Model and Hypotheses

The relationship between VA's interaction complexity and interaction outcomes

Before specifying the impact of age on VA usage, we first derive the underlying effects of interaction complexity. Based on the theoretical insights from linguistics and cognitive psychology, we assume that the number of information cues that a user must process per turn in an interaction with a VA affects how complex the user perceives the interaction to be. Accordingly, a conversation with few turns but many information cues per turn should be more complex than a conversation with many turns but few information cues per turn. This argumentation is in line with prior research on website and task complexity (Campbell 1988) and follows Berlyne's (1960) definition in which complexity increases with the number of distinguishable elements. Compared to transmitting each information cue turn by turn, memorizing multiple information cues requested by the VA in a single turn and preparing the appropriate answer requires more cognitive resources and working memory capacity, increasing users' cognitive load. Because users' cognitive load determines how complex they perceive the interaction to be (Nadkarni and Gupta 2007), we believe that users' perceived interaction complexity (PIC) increases with the number of information cues per turn. In contrast, increasing the turns and minimizing the amount of information per turn may reduce cognitive load as it minimizes the information that needs to be processed and prevents losing information that cannot be stored (Baddeley 2015). Since working memory can even be limited to one item at times (Dumas and Hartman 2008), this would ensure that cognitive overload will not occur for the low complexity condition. Thus, we propose:

H1: The number of information cues per turn increases users' perceived interaction complexity.

Based on Berlyne (1960)'s stimulus complexity theory, we propose an inverted U-shaped relationship between users' PIC and interaction outcomes. While the outcomes of an interaction with a VA can be manifold, we specifically focus on satisfaction, enjoyment, and intention to explore. We chose these outcomes because we are interested in how users evaluate their interaction with VAs, which also determines their future VA use (Moussawi et al. 2023). Since most VA use – in work-related as well as in private contexts – is likely voluntary as of now, users' evaluation of the interaction with VAs are particularly relevant for the success of VA applications.

Following the evidence from existing studies on website complexity (Geissler et al. 2001; Nadkarni and Gupta 2007), we suggest that users are most satisfied with intermediate levels of PIC. Analogously to users' interactions with a website, we expect that an interaction with a VA becomes unpleasant if the conversation seems tedious, but also if the conversation becomes overwhelming. Hence, we assume an inverted U-shaped relationship between users' PIC and satisfaction as relevant interaction outcome from extant research on website complexity (Geissler et al. 2001; Nadkarni and Gupta 2007) and in the VA context (Nagar 2023).

Furthermore, we aim to gain a better understanding of the relationship between perceived interaction complexity and affective outcomes, specifically users' enjoyment. While users' enjoyment is a significant factor of VA use experiences (Rzepka et al. 2022), affective interaction outcomes have received less attention for studies focusing on interaction complexity in general (Ghasemaghaei et al. 2019). Based on prior research, we assume that perceived interaction complexity matching users' cognitive resources should stimulate users (Guo and Poole 2009). Because being engaged in an activity increases the enjoyment of this activity (Mahnke et al. 2014), we propose that PIC increases the perceived enjoyment of an interaction with a VA as long as the task does not exceed users' cognitive resources.

Ultimately, the same mechanism likely applies to users' intention to explore additional features of the VA. We chose intention to explore as third interaction outcome because moving from simple to more complex interaction will also require the exploration of additional tasks and features of the VA. So far, VA research was focused on investigating users' initial adoption decisions. Additionally, we assume that interaction complexity is a particularly important determinant for this interaction outcome because users' intention to explore new features is dependent on the VA's usability (Motta and Quaresma 2021) and is especially relevant for older adult's use of VAs as the scenario may be unfamiliar and they may lack an overview of the VA's features (Yu et al. 2023). Thus, we propose:

H2: The relationship between interaction complexity and interaction outcomes (i.e. (a) satisfaction, (b) enjoyment, and (c) intention to explore) is characterized by an inverted U-shape.

The effect of users' age on VA's interaction complexity and interaction outcomes

Understanding the impact of users' age on the relationship between interaction complexity and interaction outcomes requires to evaluate how age affects the use of IT systems (Tams 2014; Tams 2022). Because of the ephemeral nature of spoken information, users' ability to process speech in a timely manner and before new information is presented in a sequential way is decisive for their overall interaction experience and related outcomes. Therefore, we use the theory of processing speed (Salthouse 1996) in this study to explain how users' age impacts the inverted U-shaped relationship between users' PIC and interaction outcomes. Due to reductions in brain volume and in cerebral blood flow, neural impulses slow down (Park and Festini 2017). Hence, it becomes more difficult to process and exchange information, especially when there is only limited time available. Therefore, we propose:

H3: Users' age reduces their information processing speed.

Because older adults' lower information processing speed, they are likely to reach their cognitive limitations faster than younger adults. (Salthouse 2009; Salthouse 2016). In order to exploit VA's benefits of being easy to learn and simpler than other interface types, it can be beneficial to design step-by-step interactions that focus on one piece of information at a time (Kowalski et al. 2019). Otherwise, speech interactions could exceed older users' cognitive resources. The cognitive effort of using a particular IT system can aggravate the impact of declining information processing speed (Tams 2022).Therefore, we assume that older adults will achieve optimal interaction outcomes at lower levels of PIC than younger adults, who can process information faster. Owing to their lower information processing speed, older adults cannot deal as well as younger adults with high PIC, leading to lower levels of satisfaction, enjoyment and the intention to explore for more complex interactions. This is in line with the results from Payne et al. (2011), who find that adults with lower levels of fluid abilities experience lower levels of flow compared to adults with higher fluid abilities for a given complexity. Adults with lower fluid abilities, in turn, experienced higher levels of flow for tasks with lower cognitive demands. Because the concept of flow corresponds to stimulus complexity theory, we assume a similar effect of cognitive abilities on the relationships between PIC and the interaction outcomes. In other words, we expect the maxima of the U-shaped relationships to occur for lower levels of PIC for older adults because of age-related decline in their information processing speed. Due to lower information processing speed, older adults are likely to process less information in a given period of time. Additionally, it becomes harder to manage multiple pieces of information simultaneously (Salthouse 1996).

Thus, we propose:

H4: The inverted U-shaped relationship between interaction complexity and interaction outcomes is moderated by information processing speed such that its turning point occurs at lower levels of interaction complexity when information processing speed lower and at higher levels of interaction complexity when information processing speed is higher.

Figure 2 summarizes our research model and hypotheses.

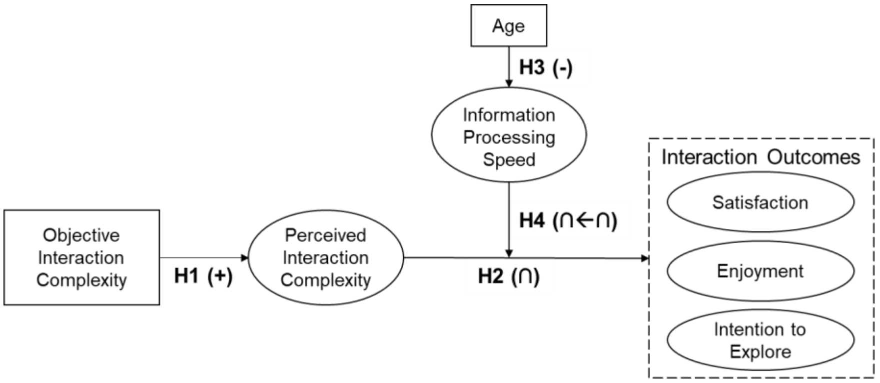


Figure 2. Research Model

Methodology

Experimental procedure

To test our research model, we designed a between-subject online experiment with four groups that interacted with VAs of varying complexity. In the experiment, participants interacted with a self-developed VA using Google's cloud platform DialogFlow, which has proven applicable in similar experiments (Diederich et al. 2020). DialogFlow is a conversational platform that uses machine learning algorithms to process natural language in human-VA interactions. We used Google's speech-to-text and text-to-speech APIs to integrate the VA in a browser interface. This interface was embedded in the survey software Qualtrics, which we used to administer the instructions and questionnaire.

The experiment proceeded as follows: After general instructions, we defined VAs and asked our participants about their prior experience with VAs. Then, we randomly assigned them to one of the experimental groups. The following interaction with our VA involved the planning of a business trip. After the participants had completed the task, we asked them to answer a questionnaire with attention checks and subjective measures of our dependent and control variables. The experiment closed with a cognitive assessment of participants' information processing speed. Figure 3 depicts the experimental procedure.

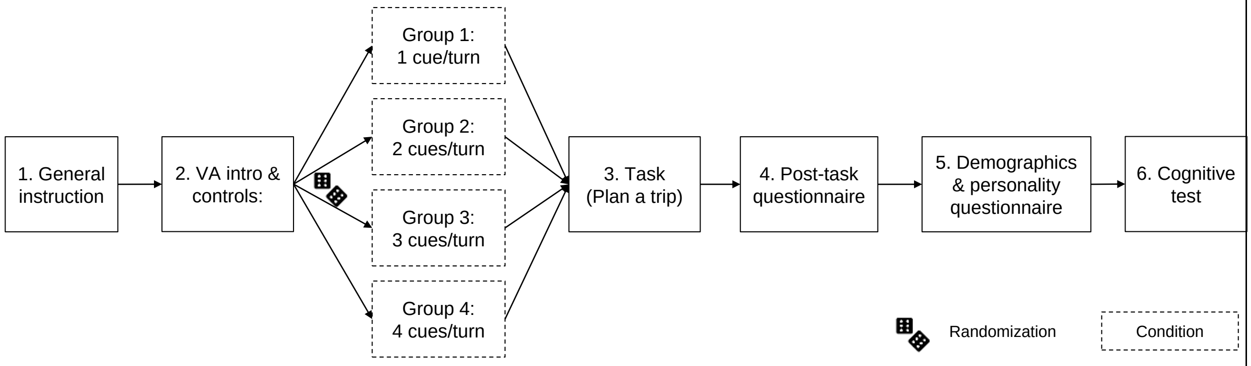


Figure 3. Experimental Procedure

We chose the travel planning task because it is a realistic VA application in both private and organizational contexts and it is relevant for both young and old adults (Georgila et al. 2010). Before the interaction with the VA, we asked participants to think of a fictitious business trip or remind themselves of a past or future trip they planned. We provided them with 12 booking parameters of a business trip (e.g., destination, departure date), which they should think about. Then, participants proceeded to the interaction with the VA, which asked them about these 12 booking parameters. Following our theoretical considerations, we manipulated the VA's level of objective interaction complexity by adapting the number of information cues which the VA requested in one turn. The number of requested information cues per turn ranged from one to four and the number of intended turns accordingly from twelve to three. We chose four information cues per turn as the most complex treatment because our working memory is able to process seven plus/minus two information cues at once (Miller 1994) and because pre-tests revealed that this number is sufficiently complex. If the VA received all 12 booking parameters from the conversation with a participant, it provided them with a password, allowing us to evaluate the completion of the task.

Measures

All latent variables' measurement scales in the questionnaire stem from existing literature to ensure content validity. We adapted the scales, if necessary, to fit the VA use context. A 7-point semantic differential scale (Bhattacherjee 2001) served as measurement of satisfaction (SAT). Moreover, 7-point Likert-type scales ranging from 'strongly disagree' (1) to 'strongly agree' (7) assessed participants' subjective perceptions of the other variables. We adapted items from Venkatesh et al. (2012) to measure perceived enjoyment (ENJ) and three items from Nambisan et al. (1999) to measure intention to explore (INT). We measured our participants' PIC of the manipulated interaction with the VA using three items of the construct 'cognitive effort' by Bulgurcu et al. (2010). For the assessment of participants' information processing speed, we used the Digit Symbol Substitution (DSS) test (Salthouse 1996; Salthouse 2016; Tams 2022). Besides measuring participants' age as a continuous variable, we accounted for further demographic variables by measuring IT self-efficacy (SEL) (Taylor and Todd 1995), trusting stance (TRU) (Mcknight et al. 2011), and personal innovativeness (PIT) (Agarwal and Prasad 1998). Right after the experimental task, a control question asked participants to enter a code that the VA provided at the end of the task. We excluded data sets from participants who failed this control question from further analyses. Additionally, we included two attention checks ("Please select 'fully agree'") to potentially exclude participants who do not read the question texts carefully. Furthermore, participants had to indicate whether they had been able to solve the task successfully or whether any technical problems had occurred. We used this question to evaluate our VA's technical performance. Finally, we asked participants for free form feedback.

We conducted two studies to test our hypotheses. Both studies were based on the identical experimental setup but differed in the sampling procedure. Study 1 tests our baseline hypotheses H1 and H2 using a balanced sample without regard to participants' age. The aim of study 1 was to evaluate whether the number of information cues per turn affects participants' PIC and to provide first evidence for an inverted U-shape relationship between PIC and interaction outcomes. Testing the baseline model in a first study before adding a moderating effect is commonly done in extant VA and elderly research (Schwede et al. 2022; Tams 2022). In study 2, we sampled participants based on their age to assess the moderating effect of their information processing speed and tested hypotheses H3 and H4.

Study 1: data collection and sample

For Study 1, we recruited a representative sample of participants that spoke English as a primary language. Participation was incentivized with a fixed participation fee for successful study completion, including passing all attention and technical checks. We required participants to use a laptop with working loudspeakers and microphone, which was tested in the beginning of the survey. Participants who did not follow our instructions were excluded from further analyses. Out of a total of 259 participants who entered the study, 138 finished the questionnaire. We excluded 14 participants who failed the technical checks and 19 who did not enter the right password after the VA interaction. Overall, we used the data from 105 participants to test hypotheses H1 and H2. The mean age was 39.13 (std. error=11.72). 49.5% of our participants in sample 1 were female. 80.0% used VAs at least weekly, while 6.7% had no experience so far.

Study 2: data collection and sample

The data collection for study 2 took place in Q3 2022. We sampled participants who were younger than 30 and older than 59 years on the survey platform Prolific. Since cognitive decline typically does not begin until individuals reach maturity in their 30s and becomes more severe around the age of 60 (Salthouse 2009), we consider these thresholds suitable for our study. We incentivized participation with a fixed participation fee and the opportunity to get an extra payment depending on the DSS score. Out of 864 participants who entered the study, 605 participants finished the questionnaire. We cleaned the data in three steps: First, we discarded four participants whose age did not fit in our desired sample. Second, we excluded 126 participants that reported technical problems and could, therefore, not solve our VA task. Third, we excluded 70 participants who did not enter the right password after the VA interaction. Overall, data from 405 participants remained for analysis. 202 participants were younger than 30, while 203 participants were older than 59. 45.2% of our participants were female. 55.8% used VAs at least weekly, while 14.8% had no experience. Table 1 depicts means and standard deviations of the variables in study 2.

	Group 1	**Group 2**	**Group 3**	**Group 4**	**All**
Age	45.02 (21.43)	46.18 (20.90)	42.94 (21.31)	43.89 (20.48)	44.63 (21.00)
DSS	25.46 (9.66)	26.42 (9.83)	28.26 (12.57)	25.22 (10.00)	26.40 (10.60)
SEL	4.95 (1.63)	4.70 (1.73)	4.95 (1.61)	4.86 (1.70)	4.85 (1.70)
TRU	4.99 (1.22)	4.90 (1.30)	5.17 (1.27)	5.10 (1.26)	5.03 (1.26)
PIT	4.57 (1.52)	4.36 (1.60)	4.20 (1.61)	4.43 (1.58)	4.38 (1.58)
PIC	3.18 (1.39)	3.46 (1.60)	4.54 (1.74)	4.64 (1.79)	3.90 (1.74)
SAT	4.83 (1.40)	4.20 (1.80)	3.51 (1.71)	3.49 (1.65)	4.03 (1.74)
ENJ	3.72 (1.72)	3.40 (1.78)	2.82 (1.55)	3.06 (1.78)	3.26 (1.74)
INT	4.17 (1.90)	3.52 (1.99)	2.98 (1.73)	3.33 (1.89)	3.50 (1.93)
N	96	125	101	83	405

Table 1. Means (and Standard Deviations)

Results

Measurement Validation

To assess whether the randomization into the four treatment groups was successful, we conducted Fisher's exact test for the categorical variables for both studies. The results confirm that there were no significant differences between participants' education, profession, and use frequency of VAs (p>.1). Furthermore, we conducted analyses of variance (ANOVAs) for the metric variables SEL, TRU, and PIT, which yielded no significant differences (p>.1).

We validated our measurement constructs PIC, SAT, ENJ, and INT as well as our control variables by exploring Cronbach's alpha (α). All constructs exhibit alpha values above .7 and are, thus, reliable. All other variables, including participants' age and DSS test scores, were assessed as single-item scores. Table 2 provides an overview of the study's key variables.

Construct	Items	Scale
PIC Study 1: α = .92 Study 2: α = .92	Using this VA is cognitively demanding.	7-point Likert-type (Bulgurcu et al. 2010)
	Using this VA requires a lot of mental effort.	
	I have to be very concentrated when I use this VA.	
SAT Study 1: α = .98 Study 2: α = .96	How do you feel about your experience of using this VA? Very dissatisfied/Very pleased	7-point semantic differential (Bhattacherjee 2001)
	Very displeased/Very pleased	
	Very frustrated/Very contented	
	Absolutely terrible/Absolutely delighted	
ENJ Study 1: α =.98 Study 2: α = .95	Using this VA is fun.	7-point Likert-type (Venkatesh et al. 2012)
	Using this VA is enjoyable.	
	Using this VA is very entertaining.	
INT Study 1: α = .99 Study 2: α = .97	I intend to explore how this VA can be used in my work tasks.	7-point Likert-type (Nambisan et al. 1999)
	I intend to explore other ways that this VA may enhance my work effectiveness.	
	I intend to spend time and effort in exploring this VA for potential applications to my work.	
DSS	The Digit Symbol Substitution (DSS) test evaluates information processing speed. The test score represents the number of correct symbols that participants can write within 90 seconds.	neuropsychological test (Tams 2022)

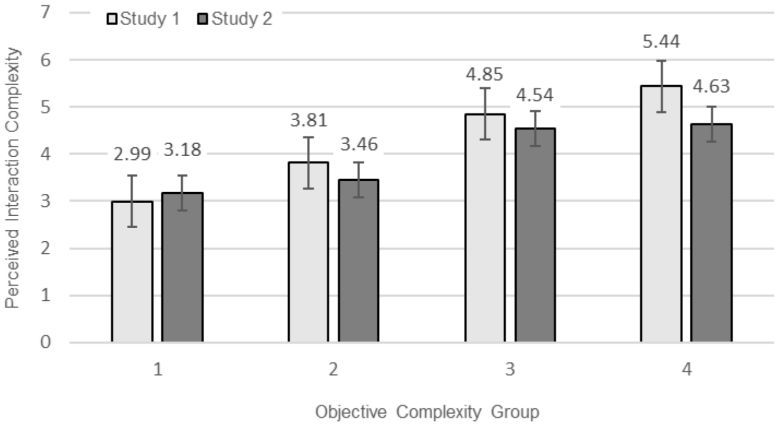
Table 2. Constructs, Items, and Scales

Hypotheses Testing

The relationship between VA's interaction complexity and interaction outcomes

We conducted an ANOVA to determine the effect of the VA's different complexity levels on participants' PIC. The results confirm significant differences between groups (Study 1: F=9.49, p<.001; Study 2: F=20.08, p<.001). Table 2 depicts the means of PIC for each experimental group and reports planned contrast between groups 1 and 2, 2 and 3, and 3 and 4.

Calculations of three planned contrasts between each group in study 1 reveal a weak significant difference between groups 1 and 2 (ΔPIC12=.83, std. error=.46, p<.1). Also, there is a statistically significant difference between groups 2 and 3 (ΔPIC23=1.04, std. error=.48, p<.05). Between groups 3 and 4, however, the difference is not significant (ΔPIC34=.59, std. error=.10, p>.1). In study 2, the difference between groups 1 and 2 is not significant (ΔPIC12=.28, std. error=.20, p>.1). The other results are consistent with study 1. The results confirm that, overall, our treatment was successful, in support of hypothesis H1. Both studies 1 and 2 show that the number of information cues per turn increases users' PIC.



Group comparisons	Contrast		Standard error		p-value	
	Study 1	Study 2	Study 1	Study 2	Study 1	Study 2
1 vs. 2	.83	.28	.46	.20	.08	.16
2 vs. 3	1.04	1.08	.48	.23	.03	.00
3 vs. 4	.59	.10	.52	.26	.26	.71

Table 2. Planned Contrasts of PIC across Treatment Groups

Following the approach for testing U-shaped relationships by Geissler et al. (2001) and Haans et al. (2016), we tested H2 by regressing our dependent variables on participants' PIC and its square, estimating the following equation: $Y = \beta_0 + \beta_1 PIC + \beta_2 PIC^2$. The results of study 1 show that for all three dependent variables, the quadratic regression model is at least weakly significant. Table 3 depicts the results of the regression analyses for study 1.

Variables	**SAT**		**ENJ**		**INT**	
Model	Step 1	Step 2	Step 1	Step 2	Step 1	Step 2
Constant	6.57*** (.38)	5.06*** (.72)	5.82*** (.41)	4.64*** (.39)	5.40*** (.46)	3.94*** (.43)
PIC	-.53*** (.08)	.43 (.40)	-.52*** (.09)	.23 (.44)	-.38*** (.10)	.55 (.48)
PIC²		-.12** (.05)		-.09* (.05)		-.11* (.06)
R²	.284	.324	.246	.268	.122	.153
Adjusted R²	.277	.311	.239	.254	.113	.137
ΔR²	.000***	.015**	.000***	.084*	.000***	.054*

Note: N=105; ***p<.01, **p<.05, *p<.1. PIC = Perceived Interaction Complexity. Standard errors in parentheses.

Table 3. Regressions of Satisfaction, Enjoyment, and Intention to Explore (Study 1)

Confirming a quadratic relationship requires further analyses (Lind and Mehlum 2010). First, β_2 should be significant at the expected sign. The significant negative coefficients for the quadratic terms indicate a inverted U-shaped relationships. Second, the slope must be sufficiently steep at both ends of the data range. We found weak support for this criterion for ENJ and INT ($p<.1$), but only partial support for SAT. For SAT, the low end of the PIC-range is positive, but not significant ($p>.1$). Third, the turning point needs to be located well within the data range. This is the case for the turning points of all three variables (SAT: 1.79, ENJ: 1.28, INT: 2.42) as the confidence interval is within the data range. Overall, we study 1 provides weak support for an inverted U-shaped relationship between PIC and the interaction outcomes ENJ and INT as hypothesized in H2b and H2c. However, the relationship between PIC and SAT in study 1 does not fulfill all criteria of a quadratic relationship.

In study 2, the quadratic regressions are also significant with a negative β_2 coefficient. However, the analyses of the slopes and turning points (SAT: 0.59, ENJ: 1.24, INT: -0.66) do not confirm the inverted U-shaped relationships. Therefore, the results of study 2 do not support H2. Since the slight differences in the results between study 1 and study 2 may be related to the differences in the samples' age distribution, we continued by investigating the moderation of these relationships through information processing speed.

The effect of users' age on VA's interaction complexity and outcomes

To test whether the DSS test captured the differences in information processing speed caused by the participants' age, we conducted a t-test. The DSS test score was significantly lower for older than for younger users in study 2 (mean difference 12.40, std. dev. .85, $p<.001$). Thus, we found support for H3.

H4 hypothesized the moderation of the inverted U-shaped relationship in terms of a shift of the vertex of the curve towards lower PIC for older users. Haans et al. (2016) recommend to test the following equation for this purpose: $Y = \beta_0 + \beta_1 PIC + \beta_2 PIC^2 + \beta_3 PIC * DSS + \beta_4 PIC^2 * DSS + \beta_5 DSS$. Table 4 depicts the results of the regression analyses. The results indicate a moderating effect of participants' DSS test score. To understand this effect in detail, we first derive the turning point by setting the first derivative with respect to PIC to zero. Then, we calculate the sign of the numerator ($\beta_1\beta_4 - \beta_2\beta_3$), which is positive for all three models, indicating that the vertex shifts to the right as the DSS score increases. Hence, with lower cognitive abilities, leading to lower DSS scores, the sweet spot in the relationships between participants' PIC and the dependent variables shifts to the left. Thus, we found support for H4.

Following up on the negative tests of the existence of an inverted U-shaped relationship between PIC and the interaction outcomes in study 2, we tested a second type of moderation: a flattening or steepening of the curve by the moderator (Haans et al. 2016). For this purpose, we evaluated the values of β_4 in each model. The regression analyses for ENJ and INT show (weakly) significant negative β_4 values, which indicates a steepening of the inverted U-shaped relationship as the DSS score increases (Haans et al. 2016). The moderation effects are illustrated in Figure 4, which depicts regression slopes for different DSS scores (low=below-median DSS score, high=above-median DSS score).

Variables	**SAT**	**ENJ**	**INT**
Constant	6.740*** (.913)	5.952*** (.981)	6.896*** (1.098)
PIC	-.827 (.540)	-.873 (.580)	-1.425** (.645)
PIC²	.020 (.068)	.045 (.073)	.114 (.082)
DSS	-.060* (.031)	-.076** (.034)	-.084** (.038)
DSS x PIC	.034* (.019)	.040* (.021)	.052** (.023)
DSS x PIC²	-.004 (.003)	-.005* (.003)	-.006** (.003)
R²	.311	.204	.189
Adjusted R²	.302	.194	.179
F Statistic	35.97***	20.39***	18.62***

Note: N=405; ***p<.01, **p<.05, *p<.1. PIC = Perceived Interaction Complexity. Standard errors in parentheses.

Table 4. Regressions of the Moderating Effect of Information Processing Speed (Study 2)

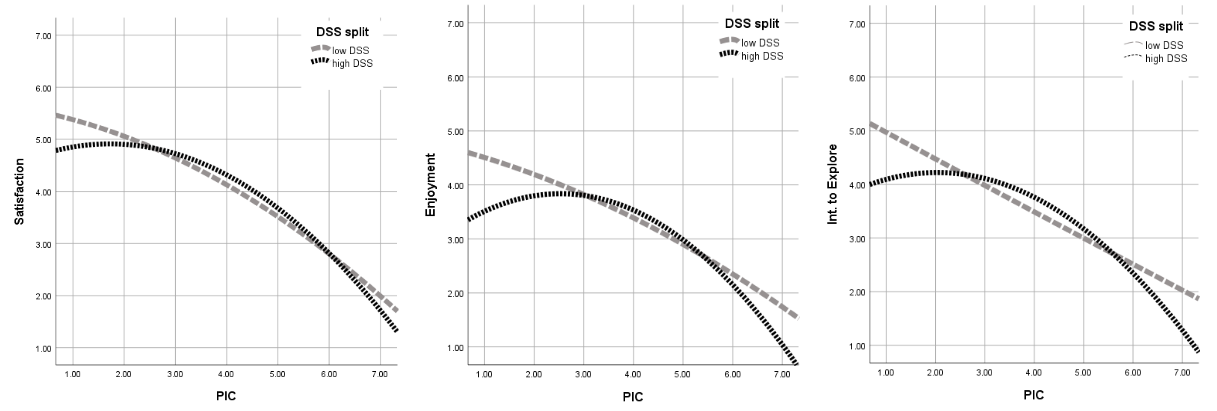


Figure 4. Regression Plots for Above and Below-Median DSS Scores (Study 2)

The figures imply that our participants' information processing speed as measured by the DSS score may not only shift the inverted U-shape but change its shape. In order to understand this effect more thoroughly, we performed a median split on participants' achieved DSS score. In cases of high DSS scores, PIC forms an inverted U-shaped relationship with the interaction outcomes SAT, ENJ, and INT. High DSS scores mean that participants' information processing speed is higher, which is usually the case for younger VA users. As the DSS score decreases, the relationship between participants' PIC and interaction outcomes becomes more linear. This means that with lower information processing speed (and higher age), VA users prefer the lowest complexity levels, compared to users with a higher DSS score, who achieve higher interaction outcomes for moderate complexity levels. Table 5 summarizes the regression results for low DSS scores, Table 6 for high DSS scores.

Variables	**SAT**		**ENJ**		**INT**	
Model	Step 1	Step 2	Step 1	Step 2	Step 1	Step 2
Constant	6.36*** (.27)	5.77*** (.55)	5.19*** (.29)	4.83 (.60)	5.59*** (.31)	5.65*** (.43)
PIC PIC²	-.60*** (.06)	-.25 (.30) -.04 (.04)	-.48*** (.07)	-.26 (.32) -.03 (.04)	-.53*** (.07)	-.57 (.35) -.00 (.04)
R² Adjusted R² ΔR²	.331 .328 .000***	.336 .329 .223	.213 .209 .000***	.215 .207 .481	.218 .214 .000***	.218 .209 .911

Note: N=405; ***p<.01, **p<.05, *p<.1. PIC = Perceived Interaction Complexity. Standard errors in parentheses.

Table 5. Regression Analyses for Satisfaction, Enjoyment, and Intention to Explore among Participants with Below-Median DSS Score (Study 2)

The results in Table 5 provide evidence that for people with lower information processing speed, only the linear regression analyses are significant, while the quadratic regressions are not. Consequently, for older adults, this means that VA interactions should be designed as easy as possible. In contrast, Table 6 offers support for an inverted U-shaped relationship for participants with higher DSS scores. All coefficients for the quadratic terms of the outcome variables are negative and significant. For low ends of the PIC-range, all slopes are positive and at least weakly significant (p<.1), while being negative and at least weakly significant (p<.1) at the high end. Furthermore, the turning points are well within the data range (SAT: 1.68, ENJ: 2.38, INT: 1.95). This means that people with higher information processing speed (usually younger) have highest subjective interaction outcomes for moderate levels of PIC.

Variables	**SAT**		**ENJ**		**INT**	
Model	Step 1	Step 2	Step 1	Step 2	Step 1	Step 2
Constant	5.84*** (.24)	4.56*** (.47)	4.59*** (.26)	3.08*** (.50)	5.00*** (.29)	3.79*** (.58)
PIC PIC^2	-.46*** (.06)	.37 (.27) -.11*** (.03)	-.35*** (.06)	.62** (.29) -.13*** (.04)	-.39*** (.07)	.39 (.24) -.10** (.04)
R^2 Adjusted R^2 ΔR^2	.237 .233 .000***	.272 .265 .002**	.140 .136 .000***	.187 .180 .001***	.131 .127 .000***	.155 .147 .017**

Note: N=405; ***p<.01, **p<.05, *p<.1. PIC = Perceived Interaction Complexity. Standard errors in parentheses.

Table 6. Regression Analyses for Satisfaction, Enjoyment, and Intention to Explore among Participants with Above-Median DSS Score (Study 2)

Discussion

In this study, we set out to investigate the relationship between VAs' interaction complexity and interaction outcomes. For this purpose, we conceptualized the objective interaction complexity for speech interactions as the amount of information requested or provided in one conversational turn. Building on this conceptualization, two research goals guided our project: First, we investigated how PIC affects interaction outcomes, specifically satisfaction, enjoyment, and intention to explore. Second, we examined the moderating effect of users' age on this relationship.

Our results confirm that users' perceived interaction complexity increases with the amount of information conveyed in one turn. This result is in line with our assumptions because each information cue requires cognitive resources, which are limited to 7 +/- cues at the same time. Similar to conceptualizations of complexity of visual interfaces (Nadkarni and Gupta 2007), VAs can provide or demand information in a more or less condensed way, balancing complexity and efficiency of information retrieval.

Additionally, we partially observed the inverted U-shaped relationship between PIC and the interaction outcomes, which we hypothesized in H2 to occur for all users. While we found general support for this relationship in an initial study with a balanced age distribution, we could not replicate the results in study 2 for both younger and older adults. Our results show, however, that the inverted U-shaped relationship is present for users with above-median information processing speed as measured by the DSS test, but not for participants with lower information processing speed. For participants with below-median information processing speed, the curve flattens to such an extent that the observed effect is rather linear than following an inverted U-shape. Analogously to the effect of website complexity (Nadkarni and Gupta 2007; Wang et al. 2014), speech interactions by younger users also benefit from a certain amount of complexity. Accordingly, speech interaction does not necessarily have to be designed as simple as possible. These results show that Berlyne's (1960) theory holds for speech interactions and, albeit only for users with high information processing speed. Among these users, however, the U-shaped effects of PIC appear for three different measures for communication effectiveness, including satisfaction, enjoyment, and intention to explore.

Furthermore, we reveal that users' age significantly moderates the relationship between PIC and interaction outcomes. Whereas we expected age to shift the vertex of the inverted U-shaped relationship towards lower levels of complexity because of declines in information processing speed, we found that it flattens the curve to such an extent that the inverted U-shaped relationship is no longer recognizable. Instead, the relationship between PIC and satisfaction, enjoyment, and intention to explore is negatively linear. Accordingly, older users tend to prefer speech interaction to be as simple as possible. While it is not surprising that older users, owing to their decline in cognitive abilities (Salthouse 2009; Salthouse 2016), shun highly complex speech interaction, we did not expect the simplest speech interaction to be the most preferred one.

Theoretical and Practical Implications

This study adds to the existing literature streams regarding interface complexity, VA usage, and elderly's use of technologies. First, we use insights form linguistic theory to extend the conceptualization of website complexity (Geissler et al. 2001; Nadkarni and Gupta 2007; Wang et al. 2014) to interaction complexity in the context of purely speech interactions. For a given quantity of information to be communicated, we show that the complexity of speech interactions depends on the number of information cues per turn. This finding helps researchers to better understand the outcomes of the interaction with VAs and to further investigate this interaction. Second, we extend previous research on the effects of IT system complexity by demonstrating that a certain amount of VA interaction complexity can be beneficial from a users' perspective. Particularly, we establish an inverted U-shaped relationship between complexity and users' satisfaction, enjoyment, and the intention to explore, given that the users have a sufficiently high information processing speed. While the literature tends to characterize the interaction with VAs as more natural and approachable for users compared to other interaction modalities, our study shows that speech interaction can become both overly and insufficiently complex. Future research should take this effect of complexity into account when designing studies on the interaction with VAs. Third, we contribute to the existing research stream on elderly's technology use and investigate how even supposedly simple and natural ways of interaction need to account for the specific characteristics of older adults. In this way, we extend previous research on the moderating effect of elderly's information processing speed (Tams 2022). Particularly, age, mediated by information processing speed, changes the nature of the relationship between interaction complexity and the interaction outcomes. Accordingly, researchers investigating the interaction with VAs must consider the age and the corresponding cognitive capacities of the population of users they are interested in.

Moreover, our findings inform practitioners on the optimal design of conversations to be neither too simple nor too complex. Since the effects of interaction complexity also depend on the individual users and their capabilities, providers of voice services should carefully consider their customer groups and how to serve them best. Providers can learn how to adapt the interaction complexity of their VAs' conversations to users' age. Specifically, our results show that interactions with younger adults with higher information processing speed levels can be designed in a more complex way. With advanced technical capabilities to infer users' age, e.g., from their speech (Burkhardt et al. 2021; Tóth et al. 2018), providers can easily adapt VAs' interaction complexity to specific user groups and, thus, enable better user experiences. In this way, they can offer less complex interactions for the elderly.

Limitations and Suggestions for Future Research

We acknowledge that our research design is not without limitations. These also open avenues for further research. First, we decided to investigate interaction complexity by manipulating the number of information cues per turn. While our results establish the interaction cues per turn as a valid instantiation of interaction complexity, this does not exclude a potential influence of other interaction design parameters such as the diversity of information per turn or the speed of talking. Future research could evaluate the main effects as well as interactions between different interaction parameters to assess their impact on VAs' interaction complexity. Another promising research avenue would be to assess the impact of varying interaction complexity levels on use-related outcomes in repeated interactions. Over multiple interactions, users may get used to a certain complexity and prefer it to simple dialogues. Second, we confronted participants with a specifically defined task to solve by interacting with a VA, setting limits to the nature and content of the conversation with the VA. While this allowed us to keep the overall amount of information to be conveyed constant, it reduces the comparability with real-world interactions with VAs. Therefore, future research may investigate how different types of tasks or the nature of the conversation change the effect of interaction complexity. Third, our study investigates the interaction with VAs in the absence of any visual cues while some modern assistance systems combine conversational and visual interaction modalities. Future research could build on our results as well as those of prior studies on the interaction with display-based systems (Geissler et al. 2001; Nadkarni and Gupta 2007; Wang et al. 2014) to evaluate the effects of complexity in multimodal interactions.

References

Agarwal, R., and Prasad, J. 1998. "A Conceptual and Operational Definition of Personal Innovativeness in the Domain of Information Technology," *Information Systems Research* (9:2), pp. 204-215.

Baddeley, A. 2015. "Working Memory in Second Language Learning," in *Working Memory in Second Language Acquisition and Processing,* Z. Wen, M.B. Mota and A. McNeill (eds.). Chicago, IL: Multilingual Matters, pp. 17-28.

Bardhan, I., Chen, H., and Karahanna, E. 2020. "Connecting Systems, Data, and People: A Multidisciplinary Research Roadmap for Chronic Disease Management," *MIS Quarterly* (44:1), pp. 185-200.

Berlyne, D.E. 1960. *Conflict, Arousal, and Curiosity*. New York, NY: McGraw-Hill.

Bhattacherjee, A. 2001. "Understanding Information Systems Continuance: An Expectation-Confirmation Model," *MIS Quarterly* (25:3), pp. 351-370.

Brewer, R., Pierce, C., Upadhyay, P., and Park, L. 2022. "An Empirical Study of Older Adult's Voice Assistant Use for Health Information Seeking," *ACM Transactions on Interactive Intelligent Systems* (12:2), Article 13.

Bulgurcu, B., Cavusoglu, H., and Benbasat, I. 2010. "Information Security Policy Compliance: An Empirical Study of Rationality-Based Beliefs and Information Security Awareness," *MIS Quarterly* (34:3), pp. 523-548.

Burkhardt, F., Brückl, M., and Schuller, B.W. 2021. "Age Classification: Comparison of Human Vs Machine Performance in Prompted and Spontaneous Speech," *Proceedings of the 32nd Conference on Electronic Speech Signal Processing*, Berlin, Germany.

Campbell, D.J. 1988. "Task Complexity: A Review and Analysis," *Academy of Management Review* (13:1), pp. 40-52.

Clark, H.H. 1996. *Using Language*. Cambridge, UK: Cambridge University Press.

Craik, F.I., and Byrd, M. 1982. "Aging and Cognitive Deficits," in *Aging and Cognitive Processes. Advances in the Study of Communication and Affect,* F.I. Craik and T. S. (eds.). Boston, MA: Springer, pp. 191-211.

Diederich, S., Brendel, A. B., Morana, S., and Kolbe, L. 2022. "On the Design of and Interaction with Conversational Agents: An Organizing and Assessing Review of Human-Computer Interaction Research," *Journal of the Association for Information Systems* (23:1), pp. 96-138.

Diederich, S., Lembcke, T.-B., Brendel, A.B., and Kolbe, L.M. 2020. "Not Human after All: Exploring the Impact of Response Failure on User Perception of Anthropomorphic Conversational Service Agents," *Proceedings of the 28th European Conference on Information Systems*, Virtual Conference.

Drew, P. 2013. "Turn Design," in *The Handbook of Conversation Analysis,* J. Sidnell and T. Stivers (eds.). West Sussex, UK: Blackwell, pp. 131-149.

Dumas, J.A., and Hartman, M. 2008. "Adult Age Differences in the Access and Deletion Functions of Inhibition," *Aging, Neuropsychology, and Cognition* (15:3), pp. 330-357.

Fisk, A.D., Czaja, S.J., Rogers, W.A., Charness, N., and Sharit, J. 2020. *Designing for Older Adults: Principles and Creative Human Factors Approaches*. Boca Raton, FL: CRC Press.

Galliers, R.D., Jarvenpaa, S.L., Chan, Y.E., and Lyytinen, K. 2012. "Strategic Information Systems: Reflections and Prospectives," *Journal of Strategic Information Systems* (2:21), pp. 85-90.

Geissler, G., Zinkhan, G., and Watson, R.T. 2001. "Web Home Page Complexity and Communication Effectiveness," *Journal of the Association for Information Systems* (2:2), pp. 1-48.

Georgila, K., Wolters, M., Moore, J.D., and Logie, R.H. 2010. "The Match Corpus: A Corpus of Older and Younger Users' Interactions with Spoken Dialogue Systems," *Language Resources and Evaluation* (44:3), pp. 221-261.

Ghasemaghaei, M., Hassanein, K., and Benbasat, I. 2019. "Assessing the Design Choices for Online Recommendation Agents for Older Adults: Older Does Not Always Mean Simpler Information Technology " *MIS Quarterly* (43:1), pp. 329-346.

Gollasch, D., and Weber, G. 2021. "Age-Related Differences in Preferences for Using Voice Assistants," in *Proceedings of Mensch Und Computer 2021*, Ingolstadt, Germany.

Guo, Y.M., and Poole, M.S. 2009. "Antecedents of Flow in Online Shopping: A Test of Alternative Models," *Information Systems Journal* (19:4), pp. 369-390.

Haans, R.F., Pieters, C., and He, Z.L. 2016. "Thinking About U: Theorizing and Testing U- and Inverted U-Shaped Relationships in Strategy Research," *Strategic Management Journal* (37:7), pp. 1177-1195.

Hall, R.H., and Hanna, P. 2004. "The Impact of Web Page Text-Background Colour Combinations on Readability, Retention, Aesthetics and Behavioural Intention," *Behaviour & Information Technology* (23:3), pp. 183-195.

Hawthorn, D. 2007. "Interface Design and Engagement with Older People," *Behaviour & Information Technology* (26:4), pp. 333-341.

Hirschberg, J., and Manning, C.D. 2015. "Advances in Natural Language Processing," *Science* (349:6245), pp. 261-266.

Hoy, M.B. 2018. "Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants," *Medical Reference Services Quarterly* (37:1), pp. 81-88.

Hsu, P.-F., Nguyen, T., Wang, C.-Y., and Huang, P.-J. 2023. "Chatbot Commerce—How Contextual Factors Affect Chatbot Effectiveness," *Electronic Markets* (33:1), Article 14.

Islam, M.U., and Chaudhry, B.M. 2023. "A Framework to Enhance User Experience of Older Adults with Speech-Based Intelligent Personal Assistants," *IEEE Access* (11), pp. 16683-16699.

Jain, M., Kumar, P., Kota, R., and Patel, S.N. 2018. "Evaluating and Informing the Design of Chatbots," *Proceedings of the Designing Interactive Systems Conference*, Hong Kong.

Kiseleva, J., Williams, K., Jiang, J., Awadallah, A.H., Crook, A.C., Zitouni, I., and Anastasakos, T. 2016. "Understanding User Satisfaction with Intelligent Assistants," *Proceedings of the 2016 ACM on Conference on Human Information Interaction and Retrieval*, Carrboro, NC.

Kock, N. 1998. "Can Communication Medium Limitations Foster Better Group Outcomes? An Action Research Study," *Information & Management* (34:5), pp. 295-305.

Kock, N. 2001. "Compensatory Adaptation to a Lean Medium: An Action Research Investigation of Electronic Communication in Process Improvement Groups," *IEEE Transactions on Professional Communication* (44:4), pp. 267-285.

Kock, N. 2004. "The Psychobiological Model: Towards a New Theory of Computer-Mediated Communication Based on Darwinian Evolution," *Organization Science* (15:3), pp. 327-348.

Kowalski, J., Jaskulska, A., Skorupska, K., Abramczuk, K., Biele, C., Kopeć, W., and Marasek, K. 2019. "Older Adults and Voice Interaction: A Pilot Study with Google Home," *Proceedings of the CHI Conference on Human Factors in Computing Systems*, Glasgow, Scotland.

Le Bigot, L., Jamet, E., and Rouet, J.-F. 2004. "Searching Information with a Natural Language Dialogue System: A Comparison of Spoken vs. Written Modalities," *Applied Ergonomics* (35:6), pp. 557-564.

Mahnke, R., Benlian, A., and Hess, T. 2014. "Flow Experience in Information Systems Research: Revisiting Its Conceptualization, Conditions, and Effects," *Proceedings of the 35th International Conference on Information Systems*, Auckland, New Zealand.

McKnight, D.H., Carter, M., Thatcher, J.B., and Clay, P.F. 2011. "Trust in a Specific Technology: An Investigation of Its Components and Measures," *ACM Transactions on Management Information Systems* (2:2), Article 12.

McTear, M. 2018. "Conversation Modelling for Chatbots: Current Approaches and Future Directions," *Proceedings of the 29th Conference on Electronic Speech Signal Processing*, Ulm, Germany.

Miller, G.A. 1994. "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information," *Psychological Review* (101:2), pp. 343-352.

Motta, I., and Quaresma, M. 2021. "Understanding Task Differences to Leverage the Usability and Adoption of Voice Assistants (VAs)," *Proceedings of the International Conference on Human-Computer Interaction.*

Moussawi, S., Koufaris, M., and Benbunan-Fich, R. 2023. "The Role of User Perceptions of Intelligence, Anthropomorphism, and Self-Extension on Continuance of Use of Personal Intelligent Agents," *European Journal of Information Systems* (32:3), pp. 601-622.

Nadkarni, S., and Gupta, R. 2007. "A Task-Based Model of Perceived Website Complexity," *MIS Quarterly* (31:3), pp. 501-524.

Nagar, K. 2023. "Customer Satisfaction with Telephone-Based Self-Service Technology: Investigating the Role of 'Gender' of the Voice Assistant," *Journal of Marketing Communications*), pp. 1-26.

Nambisan, S., Agarwal, R., and Tanniru, M. 1999. "Organizational Mechanisms for Enhancing User Innovation in Information Technology," *MIS quarterly*), pp. 365-395.

OECD. 2017. "Demographic Trends," in *Health at a Glance 2017: OECD Indicators*. Paris, France: OECD Publishing.

Oh, Y.H., Chung, K., and Ju, D.Y. 2020. "Differences in Interactions with a Conversational Agent," *International Journal of Environmental Research and Public Health* (17:9), pp. 1-20.

Palmer, J.W. 2002. "Web Site Usability, Design, and Performance Metrics," *Information Systems Research* (13:2), pp. 151-167.

Park, D.C., and Festini, S.B. 2017. "Theories of Memory and Aging: A Look at the Past and a Glimpse of the Future," *The Journals of Gerontology: Series B* (72:1), pp. 82-90.

Payne, B.R., Jackson, J.J., Noh, S.R., and Stine-Morrow, E.A. 2011. "In the Zone: Flow State and Cognition in Older Adults," *Psychology and Aging* (26:3), pp. 738-743.

Rubin, D.L., Hafer, T., and Arata, K. 2000. "Reading and Listening to Oral-Based Versus Literate-Based Discourse," *Communication Education* (49:2), pp. 121-133.

Rzepka, C., Berger, B., and Hess, T. 2022. "Voice Assistant Vs. Chatbot – Examining the Fit between Conversational Agents' Interaction Modalities and Information Search Tasks," *Information Systems Frontiers* (24:3), pp. 839-856.

Salthouse, T.A. 1996. "The Processing-Speed Theory of Adult Age Differences in Cognition," *Psychological Review* (103:3), p. 403-428.

Salthouse, T.A. 2009. "When Does Age-Related Cognitive Decline Begin?," *Neurobiology of Aging* (30:4), pp. 507-514.

Salthouse, T.A. 2016. *Theoretical Perspectives on Cognitive Aging*. New York, NY: Psychology Press.

Schick, A., Feine, J., Morana, S., Maedche, A., and Reininghaus, U. 2022. "Validity of Chatbot Use for Mental Health Assessment: Experimental Study," *JMIR Mhealth Uhealth* (10:10), e28082.

Schmitt, A., Zierau, N., Janson, A., and Leimeister, J.M. 2021. "Voice as a Contemporary Frontier of Interaction Design," *Proceedings of the 29th European Conference on Information Systems*, Virtual Conference.

Schwede, M., Zierau, N., Janson, A., Hammerschmidt, M., and Leimeister, J.M. 2022. "'I Will Follow You!'– How Recommendation Modality Impacts Processing Fluency and Purchase Intention," *Proceedings of the 43rd International Conference on Information Systems*, Copenhagen, Denmark.

Serrano, C., and Karahanna, E. 2016. "The Compensatory Interaction between User Capabilities and Technology Capabilities in Influencing Task Performance: An Empirical Assessment in Telemedicine Consultations," *MIS Quarterly* (40:3), pp. 597-621.

Tams, S. 2014. "Clashing Trends: Probing the Role of Age in Technostress," *Proceedings of the Gmunden Retreat on NeuroIS*, Gmunden, Austria.

Tams, S. 2022. "Helping Older Workers Realize Their Full Organizational Potential: A Moderated Mediation Model of Age and IT-Enabled Task Performance," *MIS Quarterly* (46:1), pp. 1-33.

Taylor, S., and Todd, P.A. 1995. "Understanding Information Technology Usage: A Test of Competing Models," *Information Systems Research* (6:2), pp. 144-176.

Tóth, L., Hoffmann, I., Gosztolya, G., Vincze, V., Szatlóczki, G., Bánréti, Z., Pákáski, M., and Kálmán, J. 2018. "A Speech Recognition-Based Solution for the Automatic Detection of Mild Cognitive Impairment from Spontaneous Speech," *Current Alzheimer Research* (15:2), pp. 130-138.

Tuohy, J.P. 2023. "Would You Let Chatgpt Control Your Smart Home?" Retrieved April 28th, 2023, from https://www.theverge.com/2023/2/24/23613382/chatgpt-smart-home-josh-ai-voice-assistant-demo.

Upadhyay, P., Heung, S., Azenkot, S., and Brewer, R.N. 2023. "Studying Exploration & Long-Term Use of Voice Assistants by Older Adults," *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, Hamburg, Germany.

Vakulenko, S., Savenkov, V., and de Rijke, M. 2020. "Conversational Browsing," *arXiv preprint arXiv:2012.03704*.

Venkatesh, V., Morris, M.G., Davis, G.B., and Davis, F.D. 2003. "User Acceptance of Information Technology: Toward a Unified View," *MIS Quarterly* (27:3), pp. 425-478.

Venkatesh, V., Thong, J.Y., and Xu, X. 2012. "Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology," *MIS Quarterly* (36:1), pp. 157-178.

Wang, Q., Yang, S., Liu, M., Cao, Z., and Ma, Q. 2014. "An Eye-Tracking Study of Website Complexity from Cognitive Load Perspective," *Decision Support Systems* (62), pp. 1-10.

Wilson, T.P., Wiemann, J.M., and Zimmerman, D.H. 1984. "Models of Turn Taking in Conversational Interaction," *Journal of Language and Social Psychology* (3:3), pp. 159-183.

Yu, J.E., Parde, N., and Chattopadhyay, D. 2023. "“Where Is History”: Toward Designing a Voice Assistant to Help Older Adults Locate Interface Features Quickly," *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, Hamburg, Germany.