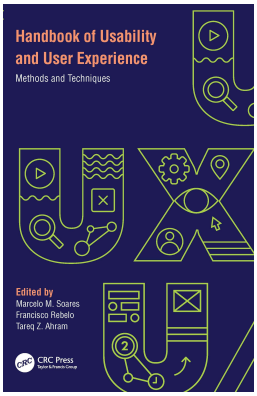

Handbook of Usability and User Experience Methods and Techniques

Marcelo M. Soares, Francisco Rebelo, Tareq Z. Ahram

Personified Virtual Assistants

Publication details
https://test.routledgehandbooks.com/doi/10.1201/9780429343490-19
Claudia Mont'Alvão, Marcela Maués
Published online on: 13 May 2022

15 Personified Virtual Assistants

Evaluating Users' Perception of Usability and UX

Claudia Mont'Alvão and Marcela Maués

CONTENTS

15.1 A BRIEF INTRODUCTION

Since the emergence of personal computers and their worldwide connection at the hand of the Internet, computing has evolved. As a consequence, the interaction with computing devices has matured as well. When there are more physical objects interconnected, and with processing capabilities of their own, the concept of computing gradually grows apart from the desktop device, leading to ubiquitous use.

Talking computers and voice user interfaces (VUI) have already been widely addressed in science fiction, but every day it is more common to find voice user interfaces on various devices such as desktop computers, smartphones, smartwatches, smart TVs and videogames. These interfaces have also been adopted in virtual

DOI: 10.1201/9780429343490-19

assistants (VAs), systems that assist their users in daily functions, ranging from organizing the agenda to controlling smart objects.

There is plenty of literature and research in the area. However, most usually address the virtual assistants' usage in English-speaking subjects, which is the standard language in most of the systems. There are not many studies addressing the virtual assistants in their translated language forms with non-English subjects. In this research case, we address the virtual assistant with Brazilian Portuguese users, which aimed to understand better if the use of voice interfaces is understood and useful in daily needs and how this happens within the singularities of this public.

15.1.1 Voice User Interfaces

In the early days of personal computing, computer interfaces would resort to metaphors within the physical world to explain the functioning and to bring familiarity to new use cases. In the current state of computing, the digital medium is already more acquainted and allows more abstract patterns for interaction.

These new interfaces should follow the same principles for visual interface design: provide feedback, navigability, consistency and attractiveness, without losing their function. In this scenario, verbal communication as an option for interface design is reinforced.

According to Nass and Brave (2007):

> Speech is a fundamental means of human communication. Even when other forms of communication—such as writing, facial expressions, or sign language—would be equally expressive, (hearing) people in all cultures persuade, inform, and build relationships primarily through speech.

These authors also affirm that in the 200,000 years of evolution, humans have become voice-activated with brains that are wired to equate voices with people and to act quickly on that identification. Talking, listening and human society have elegantly co-evolved into a remarkably interwoven, effective and stable system.

Emergent technology interfaces evolved and are now capable of producing and understanding human voices, modifying the interaction between humans and computers, proposing a new model of "natural interaction." As a result of these automatic and unconscious social responses to voice technologies, the psychology of interface speech is the psychology of human speech: voice interfaces are intrinsically social interfaces. Designers must create voice interfaces for brains that are obsessed with extracting as much social information as possible from speech and using that information to guide attitudes and behaviors (Nass and Brave, 2007). These authors also affirm that humans usually do not make distinctions among talking to a "machine/computer," or with a person, once the same part of the brain is responsible for this interaction.

Pearl (2016) determines that we are now in the "*second era of VUIs*" and the "*infancy of this next phase*." She proposes to think about users and their use cases: "*The main question to ask yourself is: will your users benefit from VUIs?*"

15.1.2 Personified Virtual Assistants (PVAs)

Weiser (1994) suggests that as new technologies and ways of interaction are offered, new systems are created as virtual assistants. They received several names as "intelligent agents," "personal virtual assistants," "intelligent personal assistants," "personal digital assistants," "mobile assistants" or "voice assistants" (McTear et al., 2016).

As examples, we can mention Siri (Apple), Google Home and Google Assistant, Cortana (Microsoft), Alexa and Echo (Amazon) and Bixby (Samsung), among many others. These virtual assistants are systems that can perform many tasks using voice interaction combined with graphics and text on the screen. McTear et al. (2016) suggest the term "conversational interface to refer to the technology that supports conversational interaction with these VPAs, by means of speech and other modalities."

As affirmed by McLean and Osei-Frimpong (2020), "in-home voice assistants are designed to be more human-like than previous attempts and intended to be an important part of an individual's everyday life." Considering the growth of voice technologies, in the early 2010s it has been verified that people are interacting with voice assistants in daily life in the same natural way as with other humans (Sundar et al., 2010).

The use of voice-control technology has become mainstream and is growing worldwide. In the United States alone, the number of people who use voice assistants—including Amazon Alexa, Apple's Siri, Google Assistant, Samsung's Bixby and Microsoft Cortana—is increasing faster than previously anticipated. It is estimated that in 2019, 111.8 million people in the United States will use a voice assistant at least monthly, up 9.5% from last year. This data is equivalent to 39.4% of Internet users and 33.8% of the total population (eMarketer, 2019). This report presents that smartphones and smart speakers are the devices of choice for most voice assistant users, and millennials are the heaviest users.

15.2 METHOD

The success of a product is not only a result of the rational characteristics of it, and therefore this research was proposed to study the relationship of users with virtual assistants. The goal of this research was to investigate acceptance and opinions about VAs and how users interact with them, considering voice as an interface. The research assumption was that the personification and the use of conversational interfaces in virtual assistants would make their acceptance easier and would make their opinions about it more positive. Some research questions conducted the literature review and the field survey:

- What is the impact of voice on the user experience of PVAs? Does the voice, instead of text, is understood as a benefit by users?
- Does the literature review reflect the particularities and field studies in Brazil? What are the similarities and points of attention about his public?
- How do Brazilians use or intend to use voice interfaces, now and in the future? What are their evaluation and opinions about PVAs?

The research was conducted in three stages. The first was the literature review that supported the discussions during the field survey among the users.

The second stage was a quantitative approach, which consisted of applying a close-ended online questionnaire aimed to understand the adoption and use of assistants in Rio de Janeiro. The questionnaire also aimed to identify participants for the qualitative phase, a focus group.

The third stage, defined after the results of the questionnaire, was a focus group technique aiming to collect data about the acceptance and the opinions of the resident of Rio de Janeiro about virtual assistants with a voice user interface. This qualitative approach was intended to map the relationship of participants with technology, highlighting its positive aspects, difficulties, concerns, as well as what they see as possibilities for future use.

15.2.1 Quantitative Phase: Online Questionnaire

15.2.1.1 Questions and Respondents' Profile

The questionnaire is the main instrument for collecting data in survey research. It is a set of standardized questions, often called items, which follow a fixed scheme to collect individual data about one or more specific topics (Lavrakas, 2008)

This technique was applied to two types of questions: open- and close-ended. The close-ended questions are those for which the answer options are predefined. The open-ended questions are those that allow the respondents to write their answers. While close-ended questions enable measurable and accurate answers, open-ended questions allow a better understanding of the respondent's thinking process, which can qualify and clarify answers.

The purpose of applying this technique in the first step of this research is to understand the adoption and the user experience of assistants in a quantitative way. It would also help to identify participants for the qualitative stage, the focus group, as previously mentioned. With the objective to collect clear and objective data in quantity, the choice was to ask close-ended questions. The only open-ended question was to request contact information when the participant told that he/she wanted to volunteer for the qualitative stage, the focus group. In this case, e-mail and telephone should be provided.

The profile was established based on the following criteria: residents of the state of Rio de Janeiro of 18 years or above. The state of Rio de Janeiro is the second Brazilian state in the use of technologies and Internet access, and therefore, significant for the gathering of such information. Also, volunteers emerged for the qualitative phase, which required a face-to-face technique. Rio de Janeiro was where the researchers were located, so it was a determinant aspect. The Consent Term was presented to all participants so that they could give their consent and proceed with the questionnaire answers.

In the questionnaire, the first questions were demographic identifiers (location, age, gender and education) and filters to narrow the search to the established profile. The ones who fit the profile determined for the survey answered the next question: if they had used any virtual assistant in the previous three months. In

that case, the following questions would be given only to those who knew voice assistants and would have a recent memory of the experience to give their opinion on it. Anyone who did not have recent contact with voice assistants would already be considered at this stage, and for the following questions, as a respondent who did not adopt this system.

Respondents who had had recent contact were conducted to answer questions regarding usage, opinion and general perceptions about the virtual assistant they had used. The questionnaire mapped which assistant was used, the access device, the purpose of use, the user evaluation of the experience and their adoption of the system. After all questions, the volunteer should state whether they would be interested in participating in the focus group—and if so—they were asked to give their contact details.

15.2.1.2 Sample and Answers

A pilot study was carried out with eight respondents, and no significant changes were made in the definitive questionnaire. Subjects were contacted using the snowball approach, considering discussion groups, professional contacts and researchers, and was answered by 200 people, of whom 161 were residents of Rio de Janeiro, older than 18 years. Among them, 102 (63.4%) had used a virtual assistant with voice interaction in the previous three months.

Their primary means of Internet access were their smartphones, laptops and desktop computers. Only ten respondents that matched the aimed profile declared smart speakers (Google Home and Amazon Echo) as Internet access devices.

The most commonly used assistants were Google Assistant, telemarketing and sales representative virtual assistants, chatbots in Web sites and apps and Apple's Siri. However, 36.6% ($n = 59$) of the respondents did not use these assistants or could not answer the question. The smartphone was the device most used, as mentioned by 93.1% ($n = 95$).

Respondents also evaluated the PVAs according to their experience and use. Experience was considered "good" by 54.7% ($n = 56$), while 20.6% ($n = 21$) considered "indifferent" and 14.7% ($n = 15$) as "bad."

Free answers in the questionnaires were valuable to define the technique for the qualitative step of the research. They highlighted different aspects of the voice as interaction. Some respondents recognized the benefits of using voice interactions, mentioning "laziness" as the main motivation.

> *Besides my experience is not good when using voice assistants, I believe that my main motivation to keep trying is laziness, mainly when I am in bed, or I do not want to look at a shiny screen. It makes me ask "Hey, Siri, what time is it?" or "Hey, Siri, how is the weather today.*" (Respondent evaluated use experience as "very bad," but keep using the Siri assistant.)

Other respondents mentioned the disadvantages of the voice as a way of interaction. One of the topics is the sharing of private information, a concern with privacy:

> *I prefer input data on the screen. I have more privacy.*

Another disadvantage is that the perception of "speed" to receive the feedback is bigger than typing on the screen, as explained by one of the participants:

> *I think that they are important (the voice assistants), but it is possible to receive the same information—and faster—with fingers.*

These answers pointed out the topics that should be investigated on a qualitative approach, considering not only the evaluation of usage and experience but also the motivations, constraints and other personal aspects that were not deeply discussed in the questionnaire.

15.2.2 Qualitative Phase: Focus Group

15.2.2.1 Technique and Its Conduction

Focus group research is one kind of qualitative research methodology. This type of research is used primarily in the social and behavioral sciences, and usually involves some interviews with people, either in groups or one-on-one. The data collected are people's views, opinions and ideas, and the data are gathered through their own words (Glitz, 1997).

The groups usually have six to ten participants. Larger groups are difficult to manage; interactions between participants tend to be less effective and discussions are hard to control. However, some techniques allow the focus group with fewer than eight participants (Krueger and Casey, 2014).

The goal is to collect data that is of interest to the researcher, which compares data among groups. Discussions must have a "focus" in a natural, logical sequence, and research aims to understand the feelings of the participants, their comments and how do they discuss the proposed topics (Krueger and Casey, 2014).

To collect Rio de Janeiro residents' acceptance and opinions about virtual assistants, as well as to map their difficulties with the technology and their expectations for the future, the focus group technique was chosen to interpret in more depth the data obtained by the questionnaire.

15.2.2.2 Pre-Session and Groups Profile

The pretest, a pilot session, was performed by a group of six males, aged 20–25, and as academic background, three were pursuing undergraduate courses, two had completed undergraduate and one was graduate. With regard to Internet access, all of them use a smartphone, desktops and laptops, and one has a smart speaker, Google Home, at home.

All of the participants were contacted a week before the meeting and instructed to ask questions on their assistants. So, all of them could participate in the debate with a recent memory for discussion. These questions were related to some answers to the questionnaire and are listed below:

- *Who are you?* Here the respondent would listen to how the personal assistant presents itself.

- *Which books Stephen King wrote?* This question mixed two languages (Brazilian Portuguese and English) in the same phrase, one of the problems mentioned by participants of the questionnaire.
- *What is the weather forecast for today?* In the questionnaire, many answers were about the use of these assistants to "make a question or search."
- *Call my mom (or dad); send a message to my boyfriend (girlfriend); how is the traffic to work/school?* These questions aimed to explore the terms used by the assistant that was not explicit (as an example, here was used "my mom" and not the mom's name). Besides, the idea was to instigate the users' curiosity about which personal information the assistant could find alone or suggest to the user.
- How do I get the PUC-Rio (university)? This question was related to the user´s geographic location in which the assistant would activate sensors and real-time information.

After checking the proposed conduction for the technique, it was verified that the tasks proposed in pre-session were adequate. Other aspects—such as the best place to set a camera or recorders; the possibility to offer some beverage and snacks—were also observed. These points were the changes to create a comfortable discussion among participants for the group sessions.

15.2.3.3 Conducting the Focus Groups Sessions

The first focus group session has six participants, one female, five males, aged 23–34, and academically they ranged from incomplete undergraduate to graduate courses. All of them used a smartphone, desktop and laptop, and just one mentioned that he did not use videogames. About PVAs, one used Cortana (Microsoft), two Siri (Apple), three Google Assistant and one never used a personal assistant, just telemarketing virtual assistants.

In the second group, seven participants attended the invitation, five females and two males, aged 20–26, and academically they ranged from incomplete to complete undergraduate courses. All of them used a smartphone, desktop and laptop, and just one mentioned using videogames. About PVAs, two used Siri (Apple) and five Google Assistant, and one never used a personal assistant, just telemarketing virtual assistants.

When each group met, the researcher conducted the session considering two discussions: first, to understand their opinion about PVA better nowadays, and second, opinions about the future.

In the first part, the researcher asked participants to redo the tasks that they did at home, once some of them said that they did not have done. After that, the groups were questioned about their impressions and opinions about these tasks.

The following topics that led the conversation in both focus groups were as follows:

- The first part—about PVAs nowadays
  Which virtual assistant have you used recently? How was your experience with this assistant? Did you like it? Why? If someone used more than

one assistant, which one you had liked most? Why? Could you perceive any similarity or difference among them?

In which language was the assistant you have used? In which device did you use it?

What was (were) your motivation(s) to use your assistant? It worked in the way you had expected? If not, tell me what happened.

Have you noticed any problem with your assistant? Have any concerns about it?

Can you remember how the voice of your assistant was? Male or female? Have you already thought about why this voice is offered in a certain gender? Did you like the voice?

What are your opinions about the assistant's "personality"? Can you remember if the PVA was sympathetic, formal, empathic, scornful, friendly? Can you imagine how this system is offered in this way? Did you like it or not?

Do you keep using the system? In each frequency. If you are not using it, tell me why.

- The second part—about PVAs in the future

Now we will talk about the future of these PVAs and your expectancies for them. Has someone here watched the movie *Her*? And the last Google presentation of Google Duplex, the system that can make an appointment at a beauty salon or restaurant? *(At this point, the researcher presented the videos mentioned to all groups, once part of the participants have not seen both videos—movie* Her *during the installation of 'Samantha' and the presentation of Google Duplex during Google IO 2018).*

What do you think about these visions for the future of PVAs? Did you like any of these assistants? Can you consider using them? To do what? How do you evaluate their utility?

What do you expect these assistants in the future? Can you imagine some future possibilities for their use? What do you want them ***to do****, ways to work? Moreover, what you* ***do not*** *want them to do?*

Do you believe that these assistants can bring problems in the future? Which kind of problems?

Do you believe that these problems can affect the adoption of PVAs by users?

And how will be the voice of these PVAs? And their personality?

Do you want to comment on something that we have already discussed or want to bring up some new points for our discussion?

A summarized result of the topics in the discussions in both focus group meetings—not considering the pretest session—is presented in the following sections, quoting non-identified phrases and words mentioned by participants as examples in italic and quotes.

15.2.2.4 Focus Group Sessions' Results and Discussion

Since the purpose of this research was exploratory, the research findings were divided into broad categories of discussions that emerged during the online form and focus groups. The results are presented under relevant topics, illustrating what the participants mentioned, and comparing them with the information obtained in the literature review.

15.2.2.4.1 Technical Limitations

This first category addresses the problems caused by the technical limitations of virtual assistants and voice interfaces. The volunteers concluded that we would overcome these technical limitations as these technologies evolve and solve the mentioned problems. Some of them also concluded that sooner or later, virtual voice assistants will end up being commonplace, because people think they are fun, and there is not a big technological gap to overcome.

- *Network and performance perception*

Voice assistants have an invocation phrase that allows them to recognize the user's intent to use them. This invocation phrase is necessary, so these systems know when the user is talking to them, and they can identify when to listen and process what has been requested actively.

Apple iOS users stated that Siri was slow to activate and to respond after the invocation and sometimes did not respond. They blame this delay after their bad smartphone performance or slow data connection and poor signal. The frustration of these users can be connected to the fact that Siri did not have clear feedback when it had no Internet access and therefore did not make it clear when it was available by the time we conducted the focus group.

> *When you say "Hey, Siri," this brief time that it takes to process the "hey Siri" and activate the feedback is too much. … You never know exactly when you could keep talking, because I cannot say "Hey, Siri, wake me up at 7 am." If I talk straight, she does not understand.*

Although complaints about performance were restricted to iOS users, users of both operating systems dislike the need to have Internet access to use the virtual assistants. They believed that Internet access should not be necessary for all the functions, and some of the tasks could be available offline.

15.2.2.4.2 Interpretation and Decision-Making

Users of both operating systems had the perception that the assistants have a limitation when it comes to listening and understanding voice. Participants argued that assistants sometimes could understand what had been said word for word, but did not know what action they should take. Nevertheless, some voice interaction problems were attributed to the assistant not understanding some words of what the user had said.

I like Siri very much, but it disturbs me when I ask her something, and the answer is something completely different (from the expected answer).

Participants stressed that assistants are more assertive when the environment is free of noise. In loud environments or when the user is in public spaces, their assistants make more mistakes. Another reported error is regarding the understanding of more than one language in the same sentence, as occurs when the user asks to play a foreign song. In the second focus group, participants stated that the assistant is better when used in English.

15.2.2.4.3 Limited Navigation

Another problem reported was the inability to use applications through speech after accessing them with the voice assistant. The assistant opens the application but does not allow continuing the interaction through the voice within it. The user must proceed through the visual interface from there. Just opening an application with a voice without the means to use it was seen as a "useless" feature.

In both focus groups, the perception that it was meaningless to use the voice when the user would have to unlock the phone and select options manually on the screen was unanimous, since "*if you are using the voice assistant, you are probably not able to use your hands*." There is an expectation of being able to usual things on the phone without manipulating anything, just interacting by voice.

One participant in the second focus group told us she did not like the way Google's approach to third-party apps within the Google Assistant. By prompting the user to talk to another entity, the user said, "*the system seems to be transferring the user to another department."* Other participants had not yet tried this functionality by the time we applied the focus group.

15.2.2.4.4 Usability

This category addresses the critical points in the use and the problems regarding effectiveness, efficiency and satisfaction—usability—of voice interfaces. These are issues that should be considered and addressed by interaction designers who are designing for this type of interface. The lack of information about the possibilities of the system leads to difficulty in use.

The most difficult assistant that I had interacted with was Cortana, at Windows. This assistant did not understand my questions, and I became confused about its functionalities, I did not understand what she could do for me.

None of the participants in either focus group could list what the assistants were able to do, and this fact was evaluated to be both "good and bad." Some participants have shown interest in exploring and discovering what their assistants could do and had fun while finding new features.

However, despite the great satisfaction when discovering something that works, a single bad experience could erode the user's confidence in that system. One of the participants exemplified this situation when he tried to impress his girlfriend, and

the system did not respond properly, culminating in frustration and abandonment of the system.

Also, some of the users have said that they do not like to explore wizards and adjust the assistant to what they know is possible and was mentioned by both Siri (Apple) and Google Assistant users:

> *After a few unsuccessful attempts, people only ask for what they know the assistant is capable of and experience less.*

Most participants in both focus groups perceived the use of voice interaction and virtual assistants as an efficient way to shorten navigation paths in the mobile devices' interfaces. Tasks that depend on many steps, such as creating an event on a calendar or calling a contact can be achieved faster. However, some participants believe that it is faster to use the visual user interface.

- *Hands-free usage*

Some of the participants in both focus groups argued that the need to unlock the phone through the screen and select options manually makes it meaningless to use the voice assistant, especially when the user cannot use their hands, such as when driving a car.

Even when hands are free to use the phone, participants mentioned that they expect that the assistant does not depend on the selection of options on the screen after a voice interaction. This case happens, for instance, in cases of need for disambiguation.

- *Invocation*

During the discussion about invocation, there were declared some preferences divergencies. While some participants said they preferred to activate the assistants by voice—affirming that press a bottom is "*boring*" once it is a "*longer interaction*"—others said that interacting with screens is better, once in this way they do not activate the assistant by accident.

This behavior matches what is related by Pearl (2016) that today a lot of us work in open space offices, and what if everybody decides to "talk" with our computer to do tasks, like "Computer, find all my Word documents this week." It will be chaos. Furthermore, "when you 'talk,' which computers are listening?"

Another point was the need to *repeat* the invocation more than once during the same conversation with the assistant. Some participants suggested that conversation must be active until the user asked to close the application or concluded the task. A counterargument was discussed that a continuous interaction could be a problem when the user is not dedicated to the interaction within the assistant.

Siri´s users, in both groups, mentioned frustration once they sometimes do not know if the system is "listening to them" when there is no signal or the signal is weak, once there is no feedback sound when the phone is soundless. Siri users agreed that

they invoke the assistant and it takes a while, or do not answer, without explaining to the users if she/he is being heard or not.

Both groups commented and agreed that inadequate feedback indicates that more development on new design options to present non-visual feedback is needed.

- *Smash the natural conversation*

Beyond the discussion about invocation, a point of complaint about all participants concerning Apple Siri was the loss of dialogue context. These users related that all phrases seem to be a "new conversation," once this assistant does not retrieve information or "remember"' of what is in the discussion. That is an attention point, the conversational interface of Apple Siri in Brazilian Portuguese is not yet well-developed and sometimes works as command interface and voice control.

Another aspect emphasized for both groups that can demotivate voice interfaces is when the assistant gives a long answer to the user, for a brief and simple question. When it happens, the user drops out a voice interface to look at the answer on the screen; in worst cases, the users feel coerced to reformulate the question, as a tentative to obtain a specific and more exact answer.

15.2.2.4.5 Accessibility

Regardless no person with special needs participate in the discussions; the participants commented that they believe that voice assistants can be useful for those who are not able to use graphic interfaces. The humanization brought by the PVA "persona" was considered positive, once the system becomes familiar. It was also mentioned in both groups that "talk" with an assistant must be faster than using a screen reader in a graphic interface.

Besides this, one of the participants criticized that the group was discussing accessibility for a chance. He does consider that:

> *"These systems were not designed to be accessible. It was just convenient that they are, once make them saleable."*

15.2.2.4.6 Other Topics Debated by the Groups

Assistants are considered as a way of entertainment, not a tool; once they are being developed and adopted more because they are fun than because they are useful.

Motivations for using the assistants are related to laziness and entertainment. They explained the laziness when they want to set a clock to wake up or when they need information and do not want to stop and look it up. Another mentioned situation was to take a shortcut during screen navigation, for example, ask for the weather forecast on a trip, search for a contact to call or set an event on the agenda.

The users consider the PVAs entertainment once they like to test the functionalities for curiosity, talk with the assistant to check joking in the way they work or to find *easter eggs* (secrets hidden in the application that has humoristic aspect).

The assuredness in both groups is that PVAs will be more useful in a fully connected world. The clear benefit is to substitute what today is done in a "mechanical way," without touching the phone, PVAs will need to know what must be bought for the pantry, control the lights, adjust the air-conditioning temperature and so on. The centerpiece was used in the domestic environment, which is noiseless, and the user has more freedom to speak with the system. However, not all of them agreed with this idea; once they remarked on it, they felt uncomfortable with the idea of controlling their lives.

- *Customization*

During discussions, participants considered as essential the PVAs adaptation to different levels and types of customization, as adequate to user´s humor, accept preferences configuration, fit the way of speaking, among other aspects.

This aspect is following what was affirmed by Nass and Brave (2007), which highlight social identity as an important point of adoption and good evaluation of voice interfaces. Dashtipour (2012) explained the Social Identity Theory (SIT):

> It has been characterized as a theory that is primarily focused on social transformation because it illustrates how social identities change and how categorization is involved in collective action. (…) Group belonging is important as a basis for self-definition. Individuals, therefore, search for positive in-group distinctiveness and discriminate against groups.

Focus group participants agreed that when everything is connected, it is easier to know everything about a person; it makes the system more personalized but implies safety and privacy negatively.

- *Humor adequacy*

Participants have the same opinion that to be "personal assistant," the system must be capable of knowing well its users and recognizing users' moods, adapting their way of communication. Users mentioned that they would not like to be treated in the same way when they are in a good mood or a bad one. One of the participants summarized it:

> *If she (Siri) is sarcastic in a day, I´m in a very bad mood; I´ll kick the phone!*

Nass and Brave (2007) also discussed mood and affirmed that users prefer consistency among the emotional state of content and voice. Users find it unusual when a sad content is said happily, and something exciting is said with a sorry voice.

- *Automation and passive customization*

In both groups, participants evaluated as positive have an assistant—in the future—that knows them very well. So well that it can solve activities without users'

interference, as setting appointments on the agenda, once it knows your schedule. In a connected world, assistants could also help to configure house appliances, like light and temperature, before the user arrives at home.

- *Configuration and active customization*

Active personalization was also a point of discussion. The PVA must adapt itself to the user but must also allow customization, considering from voice tone to system personality. This need for customization could be verified during participants' discussions: some commented that PVA could be more trickster and friendly, while others prefer a straighter and polite system. The way of interaction must be an agreement between the user and the PVA, and each person could decide whether to interact by voice or not. Just one person mentioned that it was possible to configure a male/ female voice in a PVA.

- *Filter bubble*[1]

Few participants mentioned the filter bubble during discussions, but those who mentioned demonstrated a big concern about it. To these people, users must take care with an excess of personalization, that "*can fit these PVAs in a bubble, in which you just see what you want, what you like, and what you know."* Even if it is not a problem related strictly to PVAs, once the topic was brought to the discussion, they agreed that this bubble could deepen with them, once these systems are present during the users' full day.

- *"Forced humanization" than lack of conversational interface*

Participants mentioned telemarketing virtual assistants as an example of a bad interaction, as "*a lack of natural interaction*," a "*forced humanization"* once the system speaks with you like a friend and "*keeps a long time to conclude the service*." It is related to the fact that it is not possible to "talk back" with a recorded message of these telemarketing virtual assistants. This interaction becomes artificial (or not natural) and is necessary to wait for a long set of instructions, as type an option or specific word to start the service. Participants affirmed they feel impatient during listening to all instructions, and when it is possible, they skip directly to a "real person" agent. It was also mentioned that this artificial interaction delays the service, once it is not possible to skip the preset tasks, and the "real agent" is always the last step during this interaction.

Participants considered that to simulate a real situation, the interaction must also be natural. If not, interaction is inconsistent and promotes frustration once the interaction with customer services always occurs in cases of problems, complaints and stress.

- *PVAs personality*

It is considered positive when the personalization of a PVA has a natural interaction, and it is expected that these systems can be improved in the future, according to the discussions in focus groups.

The participants that use iOs mentioned that Siri´s behavior of "*making jokes,"* "*be a rebel"* or "*talk in a sarcastic way"* gives her personality, but this aspect could be personalized by each user, as explained by one of the participants:

Siri is not ***my*** *Siri; she's de same Siri for everyone.*

During discussions, Google Assistant was considered more "neutral" than Siri and opinions were divided: for some of the users, this system is always "*passive*" and "very helpful," but "*does not have a strong personality."* For others, this assistant is "*cool*," "*friendly*" and "*helpful*."

- *"Forced" social interaction*

More introverted participants commented that they do not want to interact by voice with PVAs. They justify this decision: "*the system is one more person that I´ll need to deal with every day.*" Participants that mentioned this aspect appeared to be shy and reserved during the focus group session, and also highlighted that they are "*feeling obliged to use PVA at risk to be excluded in social groups*."

- *Social rules of voice interaction*

All participants referred to PVAs as "he" or "she," determining some human characteristics as "stupid," "polite," "friendly," "sarcastic." One of the participants said he always asks "*please"* and gives "*thanks"* to the PVA.

Even participants mentioned that they like voice interaction, some of them say they prefer text, once using during walking is not ok to speak loud, or about private issues.

This behavior is pointed out by Pearl (2016), mentioning that many people spend hours in their mobiles, usually texting. That is their default mode, and maybe they will not change to voice mode.

Even when they are not in public areas, some participants mentioned that they do not like to talk on the phone, and use chat instead, and this is the same behavior with assistants. They consider this attitude positive once they do not feel comfortable using PVAs in public, and it can be a barrier to the adoption of these systems.

As said by Pearl (2016), "Although VUIs are becoming more commonplace, not everyone feels comfortable speaking out loud to a computer, even in private."

- *Social behavior and other impacts*

Although the rich discussion about technical aspects such as usability, accessibility and customization, among others, abridged in the above topics, the major concern during discussions was the impact of PVAs on society and people's life.

a. *Voice's gender*

The fact that PVAs use as default female voices was discussed by both groups and participant's opinions were divided. For some of them, the female voice was seen as a "*heritage of subservience of female role in society"* that was transposed to PVAs. To other participants, this fact was not important once they are conditioned; once a female voice is better and preset by the system.

The female voice preference may reflect just for these discussion groups. On the other hand, the mentioned system default reinforces the stereotypes that participants consider negative for women's role in society. While it was also mentioned that both male and female voices are still "robotized voices," in the future they expect that voice scenario will become more plural.

A UN report (EQUALS, 2019) affirms that by naming voice assistants with female names (as Alexa, Cortana and Siri) and using female voices by default, tech companies have habituated users to fall back upon old-fashioned and injurious women's perceptions. The report highlights that companies were not able to design exposure controls against abusive, aggressive and gendered language. One of the problematic aspects is that the assistants have a positive reaction (as jokes) when assaulted.

b. *Privacy*

The discomfort of using PVAs in public areas was unanimous, not only to be a shy person but the "*feeling to be exposed,"* once "*there is no privacy in asking something to PVA with people observing*," even recognizing the practicality of voice interaction.

Another aspect concerning privacy is related to share personal information with strangers, not only the system. This idea led the group discussions to the need for transparency about what data is being collected by the system, how it is done, how data is stored, distributed and mainly, who will access it. Some participants said they are already "*resigned*" that their data are being stored and believe that this fact is not bad at all, once this data can personalize the system. This group defends that to design a personalized system; it must learn how the user speaks and relates with others—loose privacy is one aspect that will imply better PVAs.

All of them believed that data privacy is already a concern, and PVAs are not establishing a new problem, making the problem worse.

This aspect was also mentioned by Pearl (2016):

> If they discuss a health issue, most users won't want to do so by speaking to their phone on the train ride into work. It's not just privacy for what the user says to systems, either—it's the potential privacy violations of a VUI automatically regarding your text messages out loud."

c. *Sale and sharing of personal data*

Another concern shared by participants was the fact that companies share and use their private data. A preoccupation is that once PVAs are automatized and

can decide on the user, it gives a bias for advertisers. One of the participants gave an example of this situation if he asks the assistant to buy a pair of shoes. In this case, he believes that:

> *The system will consider not only users' preferences and specific configuration, but also will indicate a "partner" to this purchase.*

This example started a new discussion: some participants affirmed to be "invaded" with personalized ads, while others feel passive, once in this case, they received interesting ads.

They all mentioned that Web site and apps terms of service documents they need to accept or are not allowed using the product/ service, as explained by one of the participants:

> *Or you accept, or you will be excluded from everything and everyone for not using (the app), and it is human nature be included, accepted.*

d. *Reinforce of bad behavior*

A point mentioned during discussions was about personalized assistants that stimulate and reinforce ideas and bad behaviors. As an example, a racist person can personalize the assistant that will agree with anti-ethical attitudes and generate a "filter bubble."

One of the participants that has a technical profile discussed this aspect, mentioning that there is a risk not only for personalization but also for classificatory systems. In this case, the assistant could support prejudice and stereotypes when identifying user's profiles and label them. In this case, the assistant could identify, as an example, a disease in a certain user (by mistake or not) and wound her/his privacy making decisions on alerting the family or health insurance companies. All of them agreed that this kind of situation could be a disaster in people's lives.

e. *Inequality and unemployment*

Inequality was also a topic of discussion by groups. The cost of this technology is understood by participants as a segregator, once it can prevent the access of part of the population, or allow the access for just an outdated and low-quality technology. On the other hand, once technologies drive forward, they cost less and become more popular. Finally, it was also highlighted that "*in Brazil, until now (2019), there are citizens with no internet access*."

The topic of unemployment was discussed, considering how virtual assistants would substitute humans in certain jobs. Some of the participants believe that the human factor is important, and that one day he/she will lead the same situation (as customer service, for example) and will have empathy with the user. Others believe that this is naïve thinking, that one day people will have more free time in their days, that one day the technology will take humans' place in performing simple tasks.

One of the participants mentioned that this substitution of humans with machines would generate more inequality and would affect jobs needing less skill. On the other hand, even the more advanced assistants would not be able to replace professionals who perform intellectual activities, or this kind of exchange would not be financially viable.

f. *Impact on new generations and their social behavior*

During discussions it was mentioned that PVAs would leave people lazy and pampered, and they will be unprepared to perform basic tasks alone. As an example, it was cited that since smartphones became popular, people started to type and write with errors (due to the use of abbreviations), and lost the practice of handwriting.

It was also debated that virtual assistants can lead to human seclusion. Some argued that isolation would happen, once no social interaction will be necessary for daily life and in the distant future, people will create a stronger relationship with VAs rather than with other persons, as was said by one of the participants:

> *It's being designed a technology that seems a person for no need to talk with a "real" person.*

Some participants mentioned that technology would minimize daily tasks, and people will have more free time for human interactions that care. They expressed that nowadays there is an opposition between talking and interacting with people because everyone prefers using chats, and maybe virtual assistants could bring back the natural speak.

In the face of this argument, participants remembered that assistants do not have as the only function perform tasks, but also a social role. Social and amusement roles are expected with great potential for those who cannot socialize "in person," such as the aged, the ill or those with mobility problems.

15.2.2.4.7 Future Adoption

The common understanding is that virtual assistants can be useful tools, and will evolve to be in our daily lives, more than today. In the future, these assistants will be needed, and people will have a good experience with them, not only as entertainment.

15.3 TAKEAWAYS AND FUTURE RESEARCHES

The research presented in this chapter addressed the personification and the use of conversational interfaces in virtual assistants with the premise that these traits would facilitate their acceptance and make the opinion about their uses positive. At the end of the study, the result showed that the personality of these assistants indeed facilitates their adoption and has positive effects on the user's perception. However, voice user interfaces still cause discomfort when used in public environments. The biggest concerns raised by users regarding voice assistants were not technical or usability problems, but reflections on the impact of assistants on society and people's lives.

It is possible to point out some answers to the research questions:

What is the impact of voice on the user experience of PVAs? Does the voice, instead of text, is understood as a benefit by users? Besides, voice interaction was evaluated as a benefit—as the possibility of interaction with occupied hands, shortcuts to navigations, accessibility—users related that they still feel uncomfortable to use these systems in public areas, as well interact with unanimated objects (as computers and mobile phones), what can demand a difficulty in these systems adoption.

Does the literature review reflect the particularities and field studies in Brazil? Which are the similarities and points of attention about his public? Nass and Brave (2007), Pearl (2016) and UN report EQUALS (2019) are some of the authors that supported the focus groups discussion analysis, but topics like privacy, social and ethical aspects and the preference for female voices regardless of the user's gender were also observed during the discussions. Some of these points must be studied in future research.

How do Brazilians use or intend to use voice interfaces, now and in the future? How are their evaluation and opinions about PVAs? The majority of focus groups participants as questionnaire respondents does not use all the functionalities that these systems offer, but they believe that there is a potential for this technology, and consider their adoption once PVAs involute and become more usual. The connected scenario and the presence of assistants in the home environment through smart speakers are shown as potential influencers of their adoption. Results also lead to user's perception of personification and voice interface as positive aspects for adoption and experience.

Although the results can be supported by the studies conducted with the English-speaking public, there are singularities in the Portuguese-speaking public usage of PVAs.

It is not just a new technology that presents itself, but a different expectation that people have, different relations with a system that people will set in their everyday activities, with other interconnected devices, along with other human beings.

Usability and users' experience are issues that are fundamental for the adoption or not in the future and is related to decision-making, efficiency, ways of interaction (voice, text), social behavior, customization, personalization, privacy, social interaction, equality and the feeling of belonging to a society. Each of these topics needs to be investigated deeper, once they are related to cultural aspects, the role of each of us in society.

ACKNOWLEDGMENT

The authors would like to thank the Coordination of Superior Level Staff Improvement—Brazil (CAPES)—Financing Code 001 for the support for this research.

NOTE

1. "Filter bubble" is a term coined by Pariser (2011), which means an intellectual isolation that can occur when websites make use of algorithms to selectively assume the information a user would want to see, and then give information to the user according to this assumption. A filter bubble, therefore, can cause users significantly less contact with contradicting viewpoints, causing the user to become intellectually isolated.

REFERENCES

AMAZON. "What Is a Voice User Interface (VUI)?", 2019. Accessed December 29, 2019. https://developer.amazon.com/pt-br/alexa-skills-kit/vui

Dashtipour, P. *Social Identity in Question. Construction, Subjectivity and Critique*. London: Routledge, 2012. 160 p.

EMarketer. *US Voice Assistant Users 2019. Who, What, When, Where and Why*. eMarketer Report, 2019.

EQUALS. *I'd Blush If I Can: Closing Gender Divides in Digital Skills through Education*. New York: UNESCO, United Nations Educational, Scientific and Cultural Organization, 2019. Accessed December 29, 2019. https://unesdoc.unesco.org/ark:/48223/pf0000367416.

Glitz, B. "The Focus Group Technique in Library Research: An Introduction." *Bulletin of the Medical Librarian Association* 85(4), October 1997, 385–390.

Krueger, R., Casey, M. *Focus Groups: A Practical Guide for Applied Research*. 5th ed. Thousand Oaks, CA: Sage Publications, 2014. 280 p.

Lavrakas, P. *Encyclopedia of Survey Research Methods*. Thousand Oaks, CA: SAGE Publications, 2008. 1072 p.

McLean, G., Osei-Frimpong, K. "Hey Alexa… Examine the Variables Influencing the Use of Artificial Intelligent In-Home Voice Assistants." *Computers in Human Behavior* 99, 2019, 28–37. Accessed January 09, 2020. doi:10.1016/j.chb.2019.05.009.

McTear, M., Callejas, Z., Barres, D. G. "The Dawn of the Conversational Interface." In *The Conversational Interface: Talking to Smart Devices*, edited by McTear, M., Callejas, Z., Barres, D. G. New York: Springer, 2016, 11–24. doi:10.1007/978-3-319-32967-3_2.

Nass, C., Brave, S. *Wired for Speech: How Voice Activated and Advances the Human-Computer Relationship*. Cambridge, MA: MIT Press, 2007. 296 p.

Pariser, E. *The Filter Bubble: How the New Personalized Web Is Changing What We Read and How We Think*. London: Penguin Books, 2011, 308 p. Kindle format.

Pearl, C. *Designing Voice User Interfaces: Principles of Conversational Experiences*. Beijing: O'Reilly Media, 2016. 278 p.

Sundar, S. S., Xu, Q., Bellur, S. "Designing Interactivity in Media Interfaces: A Communications Perspective." In *Proceedings of the 28th International Conference on Human Factors in Computing Systems, CHI 2010: Perspectives on Design*. Atlanta, GA: ACM, April 10–15, 2010. doi:10.1145/1753326.1753666.

Weiser, M. "The World is Not a Desktop." *Interactions* 1(1), 1994, 7–8.