

This article was downloaded by: 10.2.97.136

On: 28 May 2023

Access details: *subscription number*

Publisher: *CRC Press*

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: 5 Howick Place, London SW1P 1WG, UK



## **Handbook of Usability and User Experience Research and Case Studies**

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

### **Interface Design and Usability Evaluation of Voice-Based User Interfaces**

Publication details

<https://test.routledgehandbooks.com/doi/10.1201/9780429343513-10>

Martin Maguire

**Published online on: 20 May 2022**

**How to cite :-** Martin Maguire. 20 May 2022, *Interface Design and Usability Evaluation of Voice-Based User Interfaces from: Handbook of Usability and User Experience, Research and Case Studies* CRC Press  
Accessed on: 28 May 2023

<https://test.routledgehandbooks.com/doi/10.1201/9780429343513-10>

**PLEASE SCROLL DOWN FOR DOCUMENT**

Full terms and conditions of use: <https://test.routledgehandbooks.com/legal-notices/terms>

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

---

# 7 Interface Design and Usability Evaluation of Voice-Based User Interfaces

*Martin Maguire*

## CONTENTS

7.1	Introduction .....	117
7.2	Study 1 – Survey on the Use of Voice Assistants .....	119
7.2.1	Acceptance of Voice-Based Systems.....	119
7.2.2	Use of Voice Systems and Purposes of Use .....	119
7.2.3	Functions Considered Available but Not Used .....	120
7.2.4	Usefulness of Voice-Based Assistant .....	121
7.3	Study 2 – Use of a Digital Assistant in a Student House.....	125
7.3.1	Aim .....	125
7.3.2	Method.....	126
7.3.3	Results.....	126
7.4	Study 3 – Comparison of Experienced and Inexperienced Users of a Voice-Based Digital Home Assistant .....	126
7.4.1	Aim .....	126
7.4.2	Method.....	127
7.4.3	Results.....	127
7.4.3.1	Results from Inexperienced Users .....	128
7.4.3.2	Results from Experienced Users.....	129
7.4.4	Discussion.....	131
7.4.5	Design Guidelines to Improve User Growth .....	132
7.5	Conclusion .....	133
	References.....	133

## 7.1 INTRODUCTION

Voice user interfaces allow the user to interact with a system using natural speech. Digital assistants such as Apple Siri, Google Assistant and Amazon Alexa are often used in a home setting. Voice interaction has also become a part of many consumer products in the home such as televisions, fridges and lighting. It is also available for drivers to control vehicle systems enabling them to maintain their attention on the road.

There are a number of benefits in designing a system around voice interaction:

- Data entry is possible without needing a keyboard
- It is very useful for tasks when hands and eyes are busy
- It is ideal for people who find typing difficult
- It can give direct access to command structures if statements are processed as natural language
- It is a natural means of interaction as people are good at it.

Voice interaction technology has also developed to the extent that systems can now operate fairly successfully in processing human speech without the user needing to train the system to be able to understand their voice, although errors comprehending human voice input still occur.

It is interesting then that voice interaction still seems not to be used to its full potential in everyday interactions with technology. There are some well-understood reasons for this:

- People may feel self-conscious using voice technology in a public place such as on a train or in an office alongside colleagues
- It may be distracting for others if used in a quiet area
- Speech output is transient in nature, so information becomes hard to receive and remember
- Speech output for reading text is also slower than a human scanning and reading a page of text.

It has also been found that though many “skills” or voice applications have been created for voice-based assistants, only very few have received reviews, thus indicating limited use. It has been said that with voice interfaces becoming more and more common, there is a need to make them more conversational and user-centered.

There are also well-established principles for the design of speech interfaces. Some key principles are:

- *Accommodate conversational speech:* The system should speak in a natural way and adopt human-to-human speech conventions. This acts to increase the interaction flow and comprehension (Cohen et al, 2004, and Yankelovich, et al, 1995).
- *Minimize short-term memory load:* The user’s short-term memory load should be kept to a minimum. In the absence of a companion screen display, listed information should be kept short and concise, containing only information necessary to the action being performed. The complexity of concepts the user must understand and the number of things they must learn to use the system must also be kept to a minimum (Damper and Gladstone, 2006, Kim, 2012).

- *Provide the ability to control and interrupt:* The system should allow the user to interrupt if routed to a path they do not wish to follow (Yankelovick et al., 1995, Shneiderman and Plaisant, 2010). Example: the user can either interrupt with a new interaction or simply say “stop.”

These principles have also been incorporated into a set of heuristics for speech user interactive evaluation (Maguire, 2019).

## 7.2 STUDY 1 – SURVEY ON THE USE OF VOICE ASSISTANTS

### 7.2.1 ACCEPTANCE OF VOICE-BASED SYSTEMS

A general issue relating to voice-based systems is whether consumers will accept them and how they will be used in the future. A key question is what human factor issues are preventing their use at present. A survey of 80 adults was carried out in 2019 to understand their attitudes and experiences of using voice interfaces.

The topics covered in the survey were:

- Use of voice systems and purposes of use
- Functions considered available but not used
- Usefulness of voice-based assistant
- Problems experienced in using voice-based systems
- Expectations of use in the future.

The survey received 80 responses comprising mainly young and middle-aged adults (see Table 7.1).

### 7.2.2 USE OF VOICE SYSTEMS AND PURPOSES OF USE

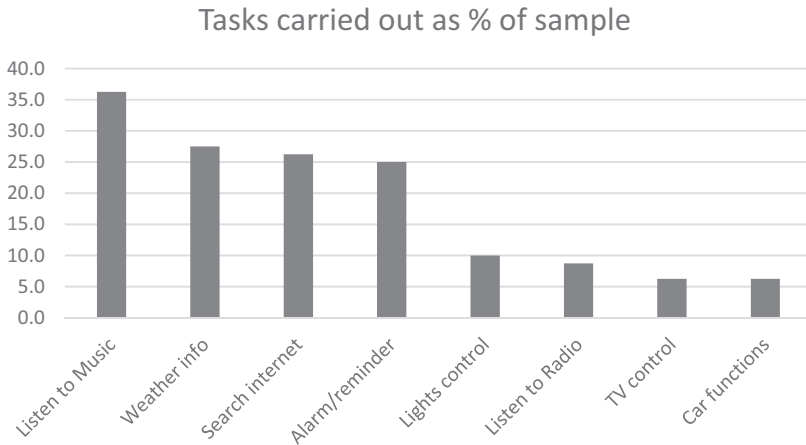
When asked which voice systems participants had experience of, the results revealed that nine different systems had been used including different types of smart speaker, PC-based systems and an in-car system. Of these instances, 96% were digital assistants (e.g., Apple Siri, Google Assistant, and Amazon Echo), while 4% were more specialist devices such as in-car voice control and automated dictation. This shows how far voice-based digital assistants have come as a mainstream consumer product.

---

**TABLE 7.1.**  
**Survey Age Group Profile**

Age group	Number responding	Percentage of sample
18–29	36	45
30–49	29	36
50–75	15	19

---



**FIGURE 7.1.** Frequency of tasks carried out with voice user interfaces.

Participants were asked to list the functions they had used their digital assistant for, including actions performed with other equipment linked to the assistant. As shown in Figure 7.1, the majority of actions were straightforward tasks that required little or no setup. They included playing music (requiring linking to own streaming service), providing weather information, searching the internet or setting a timer.

Controlling lights which required connection to the external device, i.e., the light, was carried out by 10% of the sample. As the number of devices digitally connected within the home increases, the potential for control of them through the digital assistant will increase although the setup may be a barrier to non-technical users.

Other specific tasks reported included asking to memorize friends' names; asking for directions, phone numbers and train times; receiving messages when driving; controlling AV system; dictating text to be used later; turning on torch; asking to tell a joke; adding items to the shopping list; asking it to speak Cantonese; and general banter with the device. Some tasks might be less frequent because they have not been discovered by users. Over time, the range of things people will think of that they can usefully do may increase. (There were 77 respondents to this question.)

### 7.2.3 FUNCTIONS CONSIDERED AVAILABLE BUT NOT USED

Participants were also asked to name voice-based facilities that they were aware of, but that they had not used, to get an idea of the potential growth areas for speech. Responses included more complex tasks such as buying online, controlling compatible electronics (lights, heating, CCTV, etc.), playing games, adding events to a calendar, temperature adjustment, home monitoring and helping people with disabilities. Another suggested application area was monitoring the status of elderly relatives through sensors and communicating with them in an emergency. It was thought that voice could also provide access to third-party apps on mobile devices, e.g., messaging or conferencing.

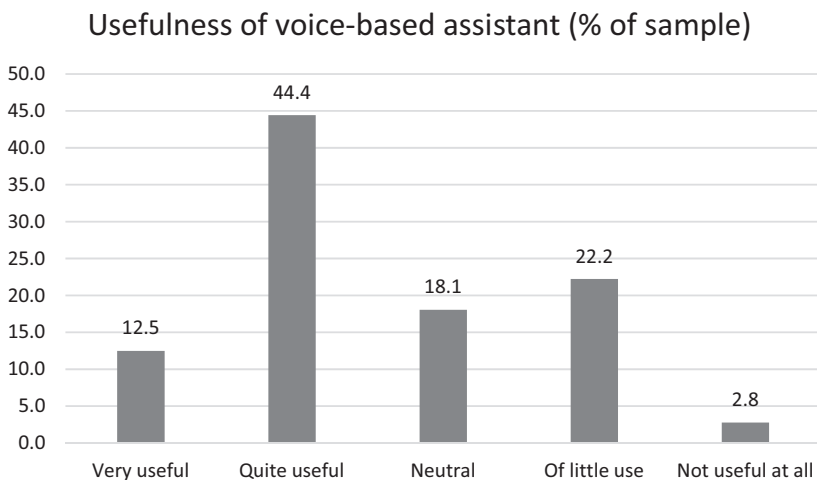
Some of these activities show the potential for more complex interactions but at the same time, requiring an information structure and feedback, which is perhaps easier to provide via a visual interface – hence the growth of smart speakers with screens. One person thought that the voice interaction could be a channel to give access to all the information that visual interaction had at the moment. This raises the interesting challenge of crafting the voice interface to work in a suitable way without simply matching the visual equivalent. (There were 61 respondents to this question.)

#### 7.2.4 USEFULNESS OF VOICE-BASED ASSISTANT

Participants also gave a rating of how useful they found their use of a voice-based assistant. A spread of responses was received with around 58% regarding it as “very useful” or “quite useful” (indicating positive support but with reservations), while one-quarter of the respondents found it of little or no use (There were 71 respondents to this question.).

Comments were also elicited to support the rating given. Around 10% of the respondents felt that using speech was useful to perform tasks when the user’s hands were not free such as when cooking, doing the laundry, driving or it was simply more convenient, e.g., to turn the lights off or the radio on when in bed, or to change music track while walking along or at a distance from the music player (Figure. 7.2).

While there was a clear view that using voice was quicker than typing for simple tasks such as internet search, playing music or setting a timer, it was felt that this was not very effective for more complex actions such as dictating documents or setting up a playlist. While it was possible to use voice for texting, using a group-messaging app such as WhatsApp was felt to be probably too hard. However, one view was that voice could be a good alternative to typing text, especially if the user is a non-native



**FIGURE 7.2.** Rating of usefulness of having voice-based assistant.

speaker. A small number of participants stated that they preferred physical button pressing, e.g., when in a car, to speech interaction, which was felt distracting.

It was said that the real benefit of speech interaction for domestic use was when devices such as lights, heating and TV are correctly connected. This can save time and mean less use of multiple control pads or remote-control devices in the home. As one respondent said, “it needs to be better incorporated into daily life.” However, it was also said that many control apps require full access to location, contacts, media, etc. “If permission is not given, then no service will be provided (it’s all or nothing).”

There was a desire for voice applications to be able to interact conversationally, but it was felt that often they were not fluent enough at present with the app misunderstanding the speaker too often. It was commented that it was interesting to see children interacting with speech technology and not knowing a world without it. (There were 70 respondents to this question.)

The survey respondents were asked to describe any problems that they had experienced in using the voice-based assistant. By far, the most common problem mentioned was a misinterpretation of the users’ input requiring them to repeat it (38%). Related to this, 8.5% of people stated that the system seemed to have problems with their local accent. Other problems that were mentioned by multiple respondents were that their speech system failed to receive inputs correctly due to interference of noise in the environment, e.g., background music, and the system responding when no question was asked. Linking up devices with the smart speaker was also mentioned. Devices can also drop off the network requiring the speaker to be rebooted. Knowing the correct commands to use for particular actions was a problem for a small number of participants, which is a feature of speech systems in contrast to screen-based systems, which can display commands. It was also felt not to be sufficiently accurate with speech-to-text applications especially putting punctuation at the right place. Table 7.2 provides a summary of the problems experienced.

These responses give an indication of the wide range of problems that can occur with speech systems, showing that it is not yet a failsafe means of interaction. (There were 71 respondents to this question.)

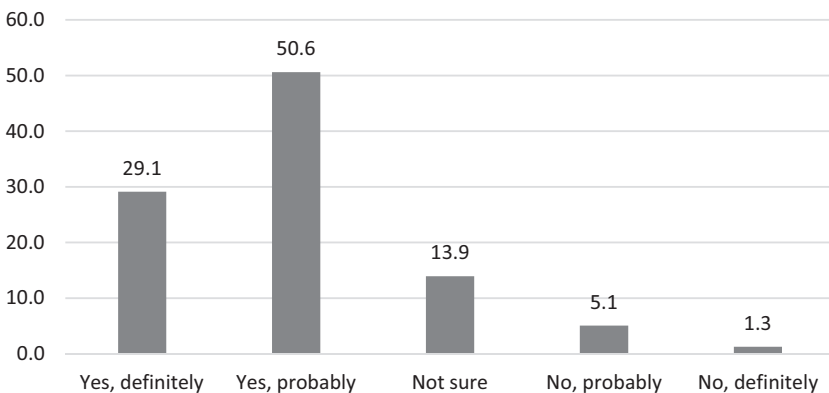
Participants were asked to indicate their opinion of whether, in the future, they think most people will be using voice-based assistants for a wide range of purposes. Figure 7.3 shows that there was a strong expectation of the widespread use of the voice-based system with nearly 80% indicating “definitely” or “probably.” Only 6.4% thought that it was unlikely to happen. These opinions are likely to be based upon people’s current experiences of using voice-based devices and the current level of recognition accuracy, which has improved dramatically over the past few years. As accuracy improves, expectations will also increase and perhaps the belief that voice can be an important means of communication in the future.

Considering age differences both with regard to usefulness ratings of speech interaction and the likelihood of future widespread use, Figure 7.4 shows the results of these questions for the three age groups 18–29, 30–49 and 50–69. The middle-aged group seemed to find voice-based user interfaces (VUIs) slightly less useful than the other two groups, while all three groups were optimistic that voice technology is likely to be an everyday technology that people use in the future.

**TABLE 7.2.**  
**Problems Experienced Using Voice User Interface**

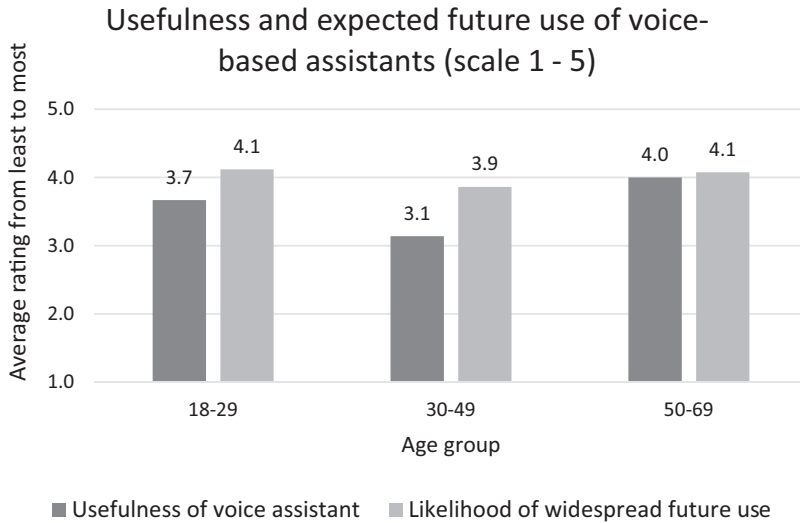
Category	Problem
Conversational	<ul style="list-style-type: none"> <li>Remembering not to say “please” at the end of an input</li> <li>Waiting for the system to finish a long answer when an input or question was misunderstood</li> <li>System not easily engaging with the user</li> <li>System requires responses too quickly before the user has thought them out</li> </ul>
Environment	<ul style="list-style-type: none"> <li>Problematic when there were multiple receivers in different rooms as the requested music may play in the wrong room</li> <li>Feeling self-conscious when using the system with others are present</li> </ul>
Accuracy and sensitivity	<ul style="list-style-type: none"> <li>Not understanding the voice of visitors</li> <li>Limitation in being able to work in multiple languages at the same time</li> <li>Needing to speak clearly to be understood</li> <li>Not understanding a music artist when their name is in non-standard spelling</li> </ul>
Performance/Usability	<ul style="list-style-type: none"> <li>Knowing the correct commands for particular actions</li> <li>Slow response times</li> <li>One system on a phone, when used in a car (in hands-free mode) needed the user to click on the screen to activate it which was not convenient</li> </ul>
Security	<ul style="list-style-type: none"> <li>One system has a “drop in” feature where it is always enabled for that person. This can cause security issues</li> </ul>
Cost	<ul style="list-style-type: none"> <li>Add on equipment to operate household devices is expensive</li> </ul>

**Expected use of voice-based assistants for wide range of purposes in the future (% of sample)**



**FIGURE 7.3.** Expected use of voice-based assistants in the future.





**FIGURE 7.4.** Comparison of age groups for usefulness of voice assistant and widespread use of voice technology.

Regarding the likeliness of future use, participants were invited to add a comment to qualify their rating. Table 7.3 shows the main themes arising from the comments. The percentage indicates the number of people in the sample who expressed a similar sentiment out of 56 who commented.

It can be seen that both positive and negative comments are listed, some more strongly positive or negative than others.

There was a general acceptance that speech technology is improving. It is becoming useful for certain tasks and will become more accepted, especially by the younger generation. Many felt that it would become integrated into our everyday lives and natural to today's children in the future. However, we may come to depend on it when it is not always the most effective way to complete tasks.

Reservations were expressed – that speech interaction will need to become secure and trustworthy and be able to cope with accents better and different voices if it is to become more accepted. Speech interaction is limited in a public or noisy environment, so other technologies such as eye-tracking, haptics and remote control through mobile devices can be the alternative. It was said that people have built up associations with certain actions being tactile such as interacting with a household device, so voice control may seem unnatural.

It was thought that voice command systems need to be more natural, smarter and responsive to the complexities of human communication. The level of intelligence that comes with voice interaction was an issue with some people who considered it to be the closest thing to a personal assistant while another view was that, while providing useful functions such as efficient internet search, they should stop “pretending to be intelligent.” The barriers to using speech are having to set up connections to

**TABLE 7.3.**  
**Comments Relating to the Potential of Speech**  
**Technology for Widespread Future Use**

Comment category	Percentage of sample
Becoming more useful, easy and quick	21.6
The technology will keep improving	14.1
It will need to be secure and trustworthy	10.7
It will become incorporated into our lives	9
It is becoming a popular means of interaction	9
I don't have a need for it	7.1
We may become over-reliant on speech	5.3
Can foresee lots of new apps arriving	3.6
I plan to try some of the new speech apps	3.6
It is the closest to being an AI assistant	3.6
It will be natural for future generations	3.6
It needs to be more accurate	3.6
Would use it more if the setup was easier	3.6
It won't be used for everything	3.6
It is impractical in some environments	3.6

other devices and the difficulties of doing this, as well as the expense of purchasing speaker-compatible devices.

Overall, the potential of speech interaction was recognized. It was said that the options available were vast although “we are still not there” yet. Knowledge about system interaction needs to improve before the use of VUI can be fully realized. It was also noted that with the rise of IoT (Internet of Things), users will want more intuitive and convenient ways to communicate with the network rather than, or as well as, touch screens or physical buttons. This analysis was based on 60 people providing a comment.

## 7.3 STUDY 2 – USE OF A DIGITAL ASSISTANT IN A STUDENT HOUSE

### 7.3.1 Aim

In this study, a smart speaker was installed into a household to see how much it was used and how attitudes changed after use (Essom, 2018). The aim was to gain an understanding of how users might interact with current voice interaction technology, and hence a field study was conducted by installing an Amazon Echo/Alexa smart speaker in the kitchen of a student house for one week. The participants were four undergraduate students between 18 and 21 years, who were aware of voice-based assistants but stated that they had limited experience of using them. The intention

was to see whether the participants used the smart speaker, for what they used it and how they felt about the technology after one week of usage.

### 7.3.2 METHOD

At the start of the week, the participants were shown how to interact with the device by using the wake-up command “Alexa” and how to issue different commands such as play music or set a timer. They could interact with Alexa during the following week as they wished. The Alexa software app was available to access the user interactions and speaker responses made during the week.

### 7.3.3 RESULTS

It was found that only 27 interactions with the speaker were made which included requests to play music or the radio, set a timer (e.g., as a wake-up alarm or for cooking purposes), ask for information or a joke, download a skill or read the news headlines. Possibly more use of the device would have occurred if it had been linked into home devices such as the control of lights or ordering groceries.

Despite the limited use, participants felt that they had interacted with the device effectively. There were comments about the limited accuracy/success rate of interactions:

Social context – feeling self-conscious when speaking to the assistant with others present and

Trust – speculation about the microphone being “always-on” and listening in to their conversations.

These results reflect the survey conducted by Milanese (2016), which showed that people’s current use of consumer voice-based assistants may be at a basic level but as more services become reliant on voice assistants and they become integrated into homes, users will become more familiar with them and less self-conscious about using them. A general lesson from the study was that users may be slow or reluctant to use a digital assistant if they are not motivated to use it in some way.

## 7.4 STUDY 3 – COMPARISON OF EXPERIENCED AND INEXPERIENCED USERS OF A VOICE-BASED DIGITAL HOME ASSISTANT

### 7.4.1 AIM

A study was carried out to explore the current use of voice-based digital home assistants and how easily they can learn new commands and their perceptions of the system after guided learning (Walker, 2019). Four participants took part in the study. They were chosen to represent different age groups, gender and levels of prior system knowledge and usage as shown in Table 7.4.

**TABLE 7.4.**  
**Participants Characteristics in Study**

Participant number	Gender	Age	Experience with digital assistants
P1	Female	22	Low
P2	Female	49	Low
P3	Male	22	High
P4	Male	52	High

**TABLE 7.5.**  
**Rating Statements on a Scale of 1 (Disagree) to 7 (Agree)**

Perceived usefulness	Perceived ease of use
It allows me to complete tasks quicker	It is easy to learn new commands
It increases my ability to complete tasks successfully	It is easily controllable
It allows me to be more productive	It is clear and understandable
It allows me to more effectively complete tasks	It is flexible in its use
It makes completing tasks easier	It is easy to become skillful
I find it useful	It is easy to use

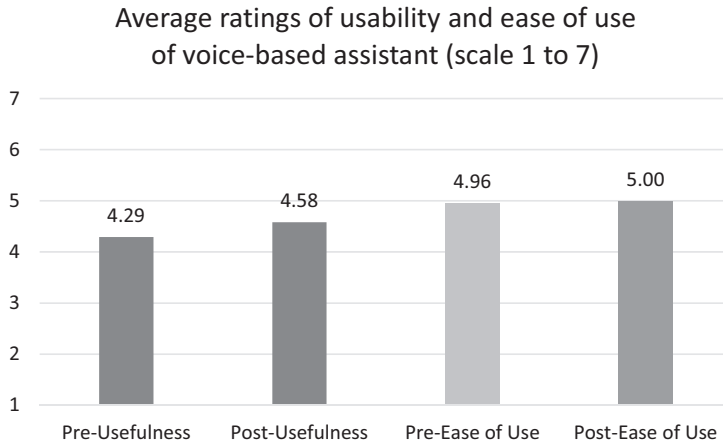
## 7.4.2 METHOD

The study took place over a period of three weeks. This comprised (1) a first week to create an understanding of each participant's regular use of digital home assistants; (2) a second week where each participant was encouraged to regularly use a selection of new commands that they were shown and (3) a third week where the participants were no longer actively encouraged to use new commands but their usage was tracked to monitor the unprompted use of the newly learned commands.

Participants were asked to complete a short questionnaire in order to gauge the perceived usefulness and ease of use of their system both before and after the study to examine whether the commands learned and experienced gained affected the way the participants perceived their system. The questions used to assess existing perceived usefulness and ease of use were based on the technology acceptance model (Davis, 1989) and adapted for use with the digital assistant. The questions are given in Table 7.5.

## 7.4.3 RESULTS

Figure 7.5 compares the mean, pre-study and post-study scores for the perceived usefulness and ease of use of the digital assistant to give an insight into the effect of the new commands on perceived usefulness and ease of use (Figure 7.5).



**FIGURE 7.5.** Comparison of participant ratings pre-study and post-study of voice-based assistant.

It can be seen average *perceived usefulness* score increased from 4.29 (pre) to 4.58 (post), an increase of  $0.29$  on the 6-point scale. The average *perceived ease of use* score changed from 4.96 (pre) to 5.00 (post), an increase of  $0.04$ . Participant's perception of usefulness increased by a modest degree after being taught new commands that they could perform with the assistant. Perception of ease of use, which was already reasonably high, only increased marginally. This indicates that being given new commands to learn was seen as helpful and did not seem to decrease ease of use.

#### 7.4.3.1 Results from Inexperienced Users

Participant P1 was a low-usage user with typical commands being “play {song name}” and “set volume to {volume number}.” They regularly used their home assistant only to play music through a linked music account and to alter the volume. P1's changes in rating from the start to the end of the study were shown in Table 7.6.

**TABLE 7.6.**  
**Ratings of Usefulness and Ease of Use for  
Participant 1 before and after the Study**

Criteria	Before	After	Change
Usefulness	4.5	4.83	+0.33
Ease of use	5.67	5.17	-0.05

The user admitted to making little to no attempt to learn or complete other tasks using the system. Due to this and the limited usage of the device, their perceived usefulness rating increased only slightly, which can be explained by always using

the system for things that they were confident in doing and a lack of desire to learn more. The ease of use score fell slightly based on ratings of “easy to learn new skills” and “easy to become skillful” being lower than pre-study. When questioned as to the lower ease of use score, P1 stated “I had to keep looking up how to complete the task ... I couldn’t find it on the app, I had to Google it.”

Participant 2 was also an inexperienced user but their usage frequency was lower, only using the system a few times a week with basic commands such as “Alexa, what’s the weather like?,” and “Alexa, how do you spell {word}.” P2 also had a more conversational style of use with questions such as “Alexa, what you up to today?” and “Alexa, what should I do today?” P2’s changes in rating from the start to the end of the study are shown in Table 7.7.

**TABLE 7.7.**  
**Ratings of Usefulness and Ease of Use for Participant 2 before and after the Study**

Criteria	Before	After	Change
Usefulness	4.17	4.50	+0.33
Ease of use	4.00	4.33	+0.33

Although the accompanying app provided with the system often attempted to suggest new commands for users to try, P2 felt that the sheer scope of possibilities hindered their willingness to learn. They compared it to being spoiled for choice as with their TV streaming service.

It was also observed that the questions and commands P2 used were similar to how one would speak to another person (whereas P1 gave commands more in line with that of speaking to machine). The phrases they used were much more diverse, with no single phrase being used regularly. Arguably, P2’s pre-study use of the system can be described as the least informed, with the lowest understanding of the potential of the system. One explanation for the large increase of 1.6 in usefulness score could be related to the user discovering the potential of the system, rather than learning new specific commands.

The main hurdle for both P1 and P2 seemed to be discovering what was possible with the speech assistant, rather than an inability to learn new commands. Both users were restricted to commands that they either already knew and were confident in using or limited by their knowledge of the style of conversation adopted with P1 communicating with the assistant as they would with a computer and P2 communicating as they would speak to another human being. This indicated a possible limitation in the way voice systems help users grow in their use of them by not making clear what was an appropriate style to communicate with them.

#### 7.4.3.2 Results from Experienced Users

Participant 3 was a more experienced younger user of the digital assistant. Their changes in rating from the start to the end of the study are shown in Table 7.8.

**TABLE 7.8.**  
**Ratings of Usefulness and Ease of Use for**  
**Participant 3 before and after the Study**

Criteria	Before	After	Change
Usefulness	4.50	4.83	+0.33
Ease of use	5.50	5.83	+0.33

P3 was the most sophisticated user from the group as they owned multiple compatible IoT (Internet of Things) products which they had linked to their digital home assistant system. These included a digital tower fan, Phillips HUE light and Sonos sound system. P3 had invested effort both in the setup of the devices and learning the commands to use them, e.g., “OK Google, turn on the fan,” “OK Google, play {song} in the kitchen,” “OK Google, turn on the lights.”

P3 had the highest score for pre-study usefulness and this can be attributed to a high level of personal investment in the system. The support to extend their use during the study led to a slightly higher usefulness rating. Despite their system being more complex, their ratings for ease of use were as high as P1, who had a simpler setup. After the study, P3’s ease of use rating was the highest of all.

It was found that P3 pre-study usage of their system was not very conversational and did not employ the day-to-day abilities that digital home assistant systems are capable of. P3 was asked to use the system as their calendar and general personal assistant for the study, using more administrative and organizational commands such as “OK Google, what do I have planned today?,” and “OK Google, can you remind me to call {name} at 6?” The aim was to encourage P3 to integrate the system into the running of their day-to-day life, not simply using it as an audio on/off button. However, it was also observed that after the study, P3 did not continue to use the new commands.

Participant 4 was a more experienced younger user of a digital home assistant. Their changes in rating from the start to the end of the study are shown in Table 7.9.

P4 used their system primarily to organize their life and conduct day-to-day activities. Examples of commands regularly used were “Alexa, call {name},” “Alexa, let {name} know I am on my way,” “Alexa, what’s in my calendar for this week?”

**TABLE 7.9.**  
**Ratings of Usefulness and Ease of Use for**  
**Participant 4 before and after the Study**

Criteria	Before	After	Change
Usefulness	4.00	4.17	+0.17
Ease of use	4.67	4.67	0

During the study, P4 was encouraged to combine all of these daily actions into a single routine command. They were shown how to create a routine in the accompanying app and then allowed to create one which best suited them. The routine that was created and used was activated by the command: “Alexa, good morning.” Using this command would trigger the digital assistant to do the following:

- Talk through the participants’ calendar and to-do list for the day
- Read out the headlines from the BBC website
- Play BBC Radio 1.

P4 used this every morning for the duration of the study and then continued to use it in the week following the study, the only participant to consistently continue to use the commands given post-study.

The results of the study showed that encouragement to learn new system commands did not increase or decrease in perceived usefulness and ease of use, in any of the four participants. So either users were satisfied with their current level of system use or the means for growing with the system were not working sufficiently well.

#### 7.4.4 DISCUSSION

The relatively small changes in usefulness and usability scores for the digital assistant show that people’s perceptions were relatively stable and were not changed much by the involvement in the study and encouragement to use new features. It could be argued that the participants in this study were close to the maximum score they could achieve in rating the voice-based digital home assistant. This could be due to users having realistic expectations pre-purchase, along with the observation that users quickly fall into a type of use; they are able to quickly learn what they find useful and then do not learn more. This raises the question of how to encourage participants to grow with the home assistant and use it in new innovative ways, and indeed whether this is desirable.

Another interesting finding was that two participants (P2 and P3) were observed to speak to their digital assistant as a machine, while the other two (P1 and P4) interacted with their systems as if speaking to a person. As digital assistants become more “intelligent,” and as users become more comfortable interacting with them, perhaps those who interact with them more like people will get more benefits and enjoyment out of them than people who simply interact with them as machines. Designers of these systems may need to encourage a new type of interaction with their systems to allow people to get the most out of them. This raises the question of how “human” the digital home assistant should be and whether it could cause a feeling of awkwardness or discomfort.

As these systems develop, become more intelligent and begin to resemble what we would more classically refer to as artificial intelligence, the way users interact with them is likely to change. At that point, further research should be conducted to determine how comfortable users find a system with a certain level of intelligence. It would theoretically allow the system to better adapt to a user’s needs but could cause



users discomfort, a phenomenon referred to as the “uncanny valley” (Pollick, 2010). Further development in making interaction more natural is to remove the need for a wake-up command and for the device to become smarter and learn spontaneous scenarios. This could be when the user utters a prominent syllable marking the start of new information. Development in this area could be for the system to recognize when new information is being delivered and when a restart may be intended for incorrect information previously stated.

#### 7.4.5 DESIGN GUIDELINES TO IMPROVE USER GROWTH

In Study 3, it was observed that without external influences users tend to quickly settle into a type of use that they rarely stray from, whether that be using it as an intelligent speaker (P1) or an audio on/off button for compatible devices (P3). These are habits seem to be based on what the individual sees as the most useful that requires the least effort. While P1 would potentially find using the system for things other than playing music useful, going through the process of identifying new commands and integrating them into regular use seems to be regarded as high effort for low perceived usefulness.

The perceived usefulness of a new command that justifies learning it is dictated primarily by the technology and what the market demands, which is out of the control of interface designers. However, what the interface can affect is the amount of effort that is required to acquire the knowledge for these new commands. Currently, the main way these systems encourage users to explore new commands and possibilities is through suggestions on the accompanying app, which often relate to a similar command that was recently used. For example, while P4 used their assistant to remind them to purchase a pack of batteries, when they next went on the app, they could be told that they could use their system to place their order through an online shop and have them delivered at their home.

Based on the results and observation of this study, the following design guidelines are suggested to decrease the perceived effort that a user needs to take to learn new commands, thus increasing the overall use of the system:

- (1) Users requirements for the system are collected during system set up so the commands prompted fit into those requirements.
- (2) Command prompts are given verbally as part of the system conversation, for example.  
User: “Alexa, set an alarm for 9 am.”  
Assistant: “Alarm set, would you also like me to read out the news?”

By moving the prompts from something that is given after use on a different medium (the accompanying application) to using them in the relevant moment, users are more likely to use the new information.

It is concluded that encouraging users to engage in using new commands to expand their use of speech systems does not seem to be an effective way to achieve greater uptake in their use. Providing guidance to use new commands at the right

time and in context could be a more useful means of user learning potential new ways to use the system and adopting them.

## 7.5 CONCLUSION

The studies report show that voice-based user interaction has a lot of potential for take-up in the home, and when given the chance to experience the capabilities of a digital assistant in the home, come to appreciate the potential that they offer. However, learnability barriers exist and by the nature of speech user interfaces, it is not as straightforward to explore possible functions as it is with visual interfaces. Even after encouraging users by suggesting new commands seemed to have limited effect unless they offered functions that they really felt they needed.

Setting up equipment to use with a digital assistant is a further challenge to the value of digital assistants unless users have experience with these technical aspects.

### ENCOURAGING LEARNING OF VOICE-BASED FUNCTIONS

One approach that can be offered to increase the use of the system is to customize the system during setup to provide the functions that the user requires and giving the assistant the capability when acting on a command and to offer further options that the user can take up.

Juang et al. (1998) outline a number of potential future developments including prosody and spontaneity. Spontaneity could be a development to remove the need for a wake-up command and for the device to become smarter and learn spontaneous scenarios. This could be when the user utters a prominent syllable marking the start of new information. Development in this area could be for the system to recognize when new information is being delivered and when a restart may be intended for incorrect information previously stated.

## REFERENCES

- Cohen, M., Giangola, J., Balogh, J. (2004) *Voice user interface design*. 1st edition. Addison-Wesley, Boston, MA.
- Damper, R., Gladstone, K. (2006) Experiences of usability evaluation of the IMAGINE voice-based interaction system. *International Journal of Speech Technology*, 9(1–2), 41–50.
- Essom, D. (2018) *The future and effectiveness of speech based automated assistants*. Design Research report, Design School, Loughborough University, Loughborough, UK.
- Kim, H.-C. (2012) An experimental study to explore usability problems of interactive voice response systems. In: Pan, J.-S., Chen, S.-M., Nguyen, N.T. (eds.), *Asian Conference on Intelligent Information and Database Systems (ACIIDS) Part III*, LNAI 7198, pp. 169–177, 201. Springer-Verlag, Berlin.
- Maguire, M. (2019) Development of a heuristic evaluation tool for voice user interfaces. In: *21st HCI International Conference Proceedings, HCII 2019*, Orlando, FL, July 26–31, eds. Aaron Marcus, and Wentao Wang, Part IV, Design, User Experience, and Usability. Practice and Case Studies. Part of the Lecture Notes in Computer Science book series (LNCS, volume 11586), Springer.

- Pollick, F. E. (2010) *In search of the uncanny valley*. Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering (LNICST), volume 40. doi: 10.1007/978-3-642-12630-7\_8
- Shneiderman, B., Plaisant, C. (2010) *Designing the user interface*. 1st edition. Addison-Wesley, Upper Saddle River, NJ.
- Walker, S. (2019) *An investigation into interfaces of artificially intelligent systems and their design*. Design Research report, Design School, Loughborough University, Loughborough, UK.
- Yankelovich, N., Levow, G., Marx, M. (1995) Designing speech acts. In: *CHI '95 Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, pp. 369–376. ACM Press/Addison-Wesley Publishing Co., New York.