



**ENHANCING USER EXPERIENCE ON SMART
MOBILE ASSISTANTS' MULTIMEDIA
NOTIFICATIONS IN APPLE SIRI & GOOGLE
ASSISTANT**

HAKAN ÖZLER

Master Thesis

Graduate School

Izmir University of Economics

İzmir

2021

**ENHANCING USER EXPERIENCE ON SMART
MOBILE ASSISTANTS' MULTIMEDIA
NOTIFICATIONS IN APPLE SIRI & GOOGLE
ASSISTANT**

HAKAN ÖZLER

A Thesis Submitted to
The Graduate School of Izmir University of Economics
Master Program in Design Studies

Izmir

2021

ABSTRACT

ENHANCING USER EXPERIENCE ON SMART MOBILE ASSISTANTS' MULTIMEDIA NOTIFICATIONS IN APPLE SIRI & GOOGLE ASSISTANT

Özler, Hakan

Master Program in Design Studies

Advisor: Assoc. Prof. Dr. Zeynep Arda

August, 2021

Interaction with a machine that has a voice was science fiction decades ago, however, it has become a part of our everyday life these days. Voice-controlled devices have been worked decades. Nevertheless, the new era was started by the launch of Apple Siri in 2011, and it is followed by Google now in the following years. The software that runs on handheld devices is described as Smart Mobile Assistant in this study. The differences and similarities among multimedia notifications, such as visual, and auditory interaction with the users and personification of Google Assistant and Siri, are the main focus of this study. Besides, human-to-smart mobile assistant (SMA) interaction regarding provided multimedia notification by the SMA, perception and reaction of the user to the multimedia notifications of the SMA, personification

tendencies of the user and its effects on the human-to-SMA interaction are evaluated in this study. Regarding personification, this study examines whether they are personified or not and what kind of narrative interweave. Moreover, future possibilities for enhancing user experience are discussed based on user research and technological opportunities. Data was collected from online interviews with content analysis. The result of this study shows that Google Assistant becomes prominent with its task-oriented features, and some sarcastic approach was indicated by the interviewed Google Assistant participants. In addition, visual notifications were found essential by some of the interviewed participants for providing visual confirmation. Additionally, it is observed that both Google Assistant and Apple Siri are personified by the users in different aspects.

Keywords: Smart mobile assistant, multimedia notification, personification, Human to computer interaction.

ÖZET

APPLE SIRI VE GOOGLE ASSISTANT ÇOKLU MEDYA BİLDİRİMLERİNDE AKILI MOBİL ASSİSTAN DENEYİMİNİN GELİŞTİRİLMESİ

Özler, Hakan

Tasarım Çalışmaları Yüksek Lisans Programı

Tez Danışmanı: Doç. Dr. Zeynep ARDA

Ağustos, 2021

Konuşan bir makine ile etkileşim yıllar önce bir bilim kurgu ögesi idi. Ancak bu günlerde günlük hayatımızın bir parçası olmaktadır. Ses kontrollü cihazlar onlarca yıldır şirketler ve laboratuvarlar tarafından çalışılmaktaydı; lakin Apple Siri'nin 2011'de piyasaya sürülmesi ve takip eden yıllarda Google Now'un rakip olarak gelmesi ile birlikte yeni bir çağ başlamış oldu. Mobil cihazlarda çalışan bu yazılımlar mobil sanal asistan (MSA) olarak bu araştırmada tanımlanmaktadır. Görsel ve işitsel bildirimler gibi çoklu medya bildirimleri arasında benzerlikler ve farklılıklar, Google Assistant'ın ve Apple Siri'nin kullanıcı ile girdiği etkileşim ve kişiselleştirilmeleri bu

araştırmanın odak noktasını oluşturmaktadır. Ek olarak MSA tarafından sağlanan çoklu medya bildirimlerinin etkileşime etkisi, kullanıcının MSA'nın sağladığı çoklu medya bildirimlerine karşı algısı ve tepkisi, kullanıcının kişiselleştirme eğilimi ve bunun insan ve MSA etkileşimine etkileri bu çalışmada incelenmektedir. Kişiselleştirme açısından bakıldığında Bu çalışma MSAların kişiselleştirilip kişiselleştirilmediğine eğer ki böyle bir durum oluşuyorsa bu örgüyü de öyküsel bir anlatımla sorgulamaktadır. Ayrıca kullanıcı deneyimini geliştirmede geleceğin getireceği olası fırsatlar da kullanıcı araştırmaları ve teknolojik fırsatlar göz önüne alınarak değerlendirilmiştir. Veriler çevrimiçi mülakatlar ile içerik analizlerinden toplanmıştır. Bu çalışmanın sonucu göstermiştir ki, Google Assistant görev odaklı özellikleri ile öne çıkmış, bazı nükteli tavırları ise katılımcılar tarafından belirtilmiştir. Ayrıca görsel bildirimler görsel teyit sağlaması açısından kullanıcıların bazıları tarafından gerekli görülmüştür. Ek olarak Hem Google Assistant'ın hem de Apple Siri'nin kullanıcılar tarafından farklı özelliklerinden dolayı kişiselleştirildikleri gözlemlenmiştir.

Anahtar Kelimeler: Mobil sanal asistan, çoklu media bildirimi, kişilik atfetme, insan-bilgisayar etkileşimi

ACKNOWLEDGEMENTS

I express my thanks to Hatun Özler and Rıza Tartan for the contribution of providing participants for the interview. Besides, I thank the interviewed participants for providing valuable information and experiences about smart assistant usage.

I would like to express my gratefulness to Assoc. Prof. Dr. Zeynep Arda for assisting to me for finding valuable resources, discovering right tools and for her supervision. Besides, I am grateful for her as advisor because of her valuable recommendations and advices.

I am very grateful to Hakan Yılmaz for providing language support and valuable resources.

Finally, I am very grateful for studying in the academy. Therefore I would like to present my gratitude to Prof. Dr. Mehmet Efe Biresseliođlu, Prof. Dr. Özgen Osman Demirbař for providing opportunities for me to work in the academy.

TABLE OF CONTENTS

ABSTRACT.....	iii
ÖZET.....	v
ACKNOWLEDGEMENTS.....	vii
TABLE OF CONTENTS.....	viii
LIST OF FIGURES.....	ix
CHAPTER 1: INTRODUCTION.....	1
1.1. Hypothesis / Design Problem.....	1
1.2. Methodology.....	2
CHAPTER 2: HISTORY OF SMART MOBILE ASSISTANT.....	4
CHAPTER 3: PRELIMINARY FINDINGS OF THE USER RESEARCH.....	8
CHAPTER 4: EVALUATING USER EXPERIENCE ON SMA INTERACTION...	18
4.1. Demonstration of Differences and Similarities Between the Notifications of Apple Siri and Google Assistant.....	18
4.1.1. Evaluating Variety of Multimedia Feedback of SMA During Interactions to User.....	18
4.1.2. Language Support.....	22
4.1.3. Analysis of User Reaction to Multimedia Notification of SMA.....	23
4.2. A Study on the Possibilities of Improving User Experience through SMAs' Multimedia Feedback.....	25
4.3. Effects of Personification on User Experience	31
4.3.1. A Study on the Personification Ability of Human Psychology: Heritage or Imprecation.....	31
4.3.2. Determining Differences of SMA with Considering Personification Side of Smart Mobile Assistants.....	39
CHAPTER 5: DISCUSSION.....	43
5.1. Limitations.....	45
CHAPTER 6. FUTURE DESIGN OPPORTUNITIES.....	46
CHAPTER 7. CONCLUSION.....	51
REFERENCES.....	55

LIST OF FIGURES

Figure 1. Preferred mobile operation system by user.....	8
Figure 2. Average experience of users on smart mobile assistant.....	8
Figure 3. Satisfaction rate of SMA users. Zero is the lowest, five is the highest.....	9
Figure 4. Location preference where SMA is accessed by user for assist.....	9
Figure 5. Frequency of visual notifications versus auditory notifications.....	10
Figure 6. Text and auditory based notifications that is generally accepted as sufficient by SMA users.....	10
Figure 7. Gratification rate of SMA voice by user.....	11
Figure 8. Evaluation of SMA voice by user as human-like or machine-like.....	11
Figure 9. Gender of SMA voice, which is preferred by SMA users.....	11
Figure 10. Emoji usage tendency of SMA which is reported by user.....	12
Figure 11. According to user opinion, congruity of SMA emoji usage.....	13
Figure 12. During interaction, emojis are frequently used by Google Assistant.....	14
Figure 13. Hints that are displayed on Google Assistant interface when it is launched.....	15
Figure 14. Provided visual and auditory confirmation notifications by Google Assistant and Siri.....	19
Figure 15. Example of two leading displays with the same speed but different jump length settings.....	21
Figure 16. Functional analysis of user evaluation data.....	22
Figure 17. Siri has relatively superiority in terms of humour. However, Google Assistant has a limited and different approach to humour.....	34
Figure 18. The table shows preferred chatbots. Sarah is mostly preferred one.....	37
Figure 19. Rated current and expected capabilities of SMA by the frequent and infrequent users.....	37
Figure 20. More vivid interaction is offered by Google Assistant. These figures were prepared for demonstrating richer multimedia interaction opportunities for Siri.....	44
Figure 21. These figures were prepared for demonstrating responses of SMA that estimates different types of users' requirements and needs.....	49

Figure 22. Google Assistant is evaluated as ice-cold-librarian. these figures were prepared based on the user expectations.....51



CHAPTER 1: INTRODUCTION

1.1 Hypothesis / Design Problem

People have desired to talk to computers since their invention. Talking computer notion have been depicted in science fiction, such as Starship Enterprise's computer talking computer in KITT car, android or robots, which are named "3-CPO" and "R2D2" in the Star Wars series. In appreciation of technology, interaction with a computer that has a voice was science fiction decades ago, nevertheless, it becomes real today. By the way, the contribution of fast internet connection is considerable. Beyond internet speed, the development of calculation power of CPU is remarkable. For instance, the power of the Apollo 11 computer processor was 0.043 Mhz. Today, the typical smartphone has a hundred thousand times more powerful processor than the Apollo 11 computer had. In spite of this immense amount of power engineers still have issues about running Smart Mobile Assistants solely by local processor and storage. There are some attempts to run SMAs in offline mode. But they still are dependent on internet connection. A smart mobile assistant is a software agent that runs on mobile devices (Hoy, 2018). When the software receives a wake-up command, it is activated. It receives commands of users and dictates texts and sends them to the specialized server. The server analyses this command, subsequently generates possible responses and actions, then it is sent to terminal devices software to perform the required task. The first smart mobile assistant, named "Siri" was launched by Apple Inc. in 2011. It has been followed by a rival called "Google Now", which was released by Google LLC. one year after launching Siri. Later, Google Now evolved into today's "Google Assistant". Even though many smart mobile assistants have entered the market, Apple's Siri and Google Assistant are still dominating it.

In this research, Apple Siri and Google Assistant are evaluated with a special focus on the feedback provided by SMAs: How they interact with users, and their personification. In terms of personification, this research tries to determine whether they are personified or not, and if so, what kind of narrative they interweave."

According to user interviews, the most complaining issue was the interruption of hands-free interaction. SMAs are mostly preferred by users while their hands are occupied by another job. Hence, interrupting the mentioned job at hand is a frustrating

experience. Besides, it contradicts the user's mental model. Similar to interhuman interaction (in which touch-based interaction is rarely used), SMA should be able to provide a bypass method of touch-based interaction by means of providing voice-only commands for the users. Secondly, some limitations are mentioned during interviews; For example, showing links that demonstrates how a job is accomplished is heavily criticized by one participant. Moreover, customization is another complaint issue on SMA interaction. Struggling with teaching her husband's name to SMA is heavily criticized by one participant. Besides, it negatively affects the personification of SMA. Because the SMAs are expected to know their users' needs, desires and behaviours. Even though multimedia notifications are found sufficient in terms of abundance and correctness of combining each other, it is not evaluated as necessary or noticeable by some participants, especially by participants who prefer SMA in driving.

Distinctly, there are many differences in the visual interfaces of Google Assistant and Apple Siri. Beyond text and background colour differences, many disparities in emoji usage can be observed. Emojis and stickers are preferred by Google Assistant. these are not used by Apple Siri. The chatbox interface of Google Assistant resembles the classical chatbox. Besides, the sensation of chatting with a real person is indicated by one Google Assistant user. Besides, using emojis on the chat interface leads to Google Assistant being personified, which is observed during interviews. According to interviews, Google Assistant tends to be personified more than Apple Siri. Additionally, a marriage proposal was declared to Apple Siri by one depressive participant. Another effect that causes personification is the humour abilities of SMA. Most of the interviews and conducted literature review agree that Siri's humour ability and sarcastic personality are indicated by many participants. Nevertheless, even though Google Assistant is depicted as a task-oriented and cool librarian resemblance personality. Some senses of humour and a joyful approach were reported by some interviewed participants.

1.2. Methodology

While personification of SMA is briefly investigated in the existing literature, multimedia notifications in the SMA experience are barely evaluated. Considering technological developments and the latest updates, the user experience has evolved

over time. In order to reach the final and investigated information, the research focuses mainly on the following questions:

1. Varieties of multimedia notifications in SMAs and the user's perception of them.
2. Relationship between the multimedia notifications and personification tendencies on the SMAs as defined by the user.

The focus group for this study was composed of ten people, which includes three female and seven male participants. The age range of the participants was quite wide, expanding from 20 to 50. The average age of participants was 37 years. The interviewed participants have been nominated as "P" -that is the abbreviation of participant- followed by a number, which is dedicated to the participant and preferred Smart Assistant by the participant. "P1 A Siri user" is given as an example. For reaching a detailed user experience, face-to-face interview was chosen as the data collection method. Because of the pandemic, the interviews were conducted online. Besides, as an unexpected fortunate outcome, social effects during the interview and unwanted psychological pressure on participants were reduced by the use of this method. With the help of online interviews, user facial expressions, tone of voice, and feelings were gathered and recorded.

CHAPTER 2: HISTORY OF SMART MOBILE ASSISTANT

The smart mobile assistants have increased in popularity since Apple introduced one in 2011. However, they go way back to before Apple launched Siri. The first ancestors of smart mobile assistants were mostly voice-activated devices. The toy dog Radio Rex, which was released in 1910, was the first voice-activated device. Upon hearing its name, it would go out of its doghouse (Krazit, 2010). These technological developments were not restricted to just walking, moving, or lighting a lamp. In 1939, Bell laboratories exhibited “Voder” at the New York World’s Fair. The name ‘Voder’ comes from ‘voice’ and ‘demonstrator’. Voder imitated human speech, not through recorded voice or human vocal cords. Voder’s voice was generated by an electronic circuit (Pieraccini, 2012). Thirteen years after the development of Voder, three scientists from Bell Labs (K. H. Davis, Rulon Biddulph, and Stephen Balashek) published a detailed description of a machine, which recognizes spoken digits. Even though it had serious problems recognizing words, Audrey (Automatic Digit Recognition) was the first speech recognition device in the world (Pieraccini, 2012). In 1962, visitors of the World’s Fair in Seattle witnessed a pioneering device in speech recognition technology. IBM Shoebox was able to recognize and respond to 16 spoken words. These included ten digits from 0 to 9, as well as command words such as “add”, “subtract” etc... Shoebox is assumed to be the ancestor of today’s speech recognition system (Time, 1961). In 1976, the Harpy Speech Recognition System was introduced as a more evolved version of the two-speech system Hearsay-I and Dragon. The Harpy System aimed to combine the best features of the two systems under the Harpy structure. Developers gained vast improvement on the accuracy of Harpy. The word recognition accuracy of this system was between 90.8% and 97.5%. It was a significant step towards overcoming the problem of pronunciation (Lowerre, 1976). Nuance Communications released the Dragon Dictate system in 1990. It was a macOS application that enabled users to interact with the computer through their voices. Users were able to do the following on their computers via Dragon Dictate:

- Dictating texts
- Editing texts using verbal commands
- Dictating application commands in order to control computer operations
- Manipulating pointers using verbal commands. (Langer, 2012)

The Defense Advanced Research Project Agency (DARPA) was working on a project called the Personalized Assistant that Learns (PAL). This project aimed to develop a cognitive computing system that makes military decision-making more efficient and effective. Moreover, another aspect of DARPA's PAL program was enabling voice-based interaction between civilian users and their handheld devices. This feature of the PAL program led to the launch of Siri Inc., which was later purchased by Apple Inc. in 2011. Thus, the first Smart Mobile Assistant appeared on the market as the Siri application for iOS (Dillow, 2011).

On the other hand, the monopoly of Siri lasted until 2012. Google released Google Now in 2012, which then evolved into Google Assistant. Microsoft released Cortana in 2013, followed by Amazon Alexa and Amazon Echo entering the market in 2014. Today, there is an abundance of smart assistants. Amazon mostly focuses on home smart speakers. As a result, it has considerable domination in the home speaker market. Nevertheless, Siri and Google Assistant have predominance in terms of smart mobile assistants in handheld devices.

In this study, the smart mobile assistant is analyzed for the following reasons. First of all, only a small portion of homes have smart speakers. Customers have to pay extra money to purchase a home speaker. Although if the customer has a smart device, it may act as a virtual assistant already. So, it makes the smart mobile assistant market far larger than the home speaker market. Secondly, smart home speaker markets are largely dominated by Amazon. Nevertheless, Apple Siri and Google Assistant mostly share the smart mobile assistant market because of their success on the handheld devices. Third, multilingualism has positive effects on customer's decisions. For example, while Google Assistant supports forty-four languages, Alexa just covers eight languages (Templeton, 2020). Lastly, a Smart mobile assistant interacts with users through various media outputs, such as video, image, and sound. In addition, this research aims to evaluate smart mobile assistants in terms of multimedia interaction. Moreover, home assistants suffer from a lack of multimedia feedback opportunities.

Interaction design played an essential role in human-to-SMA interaction. Interaction design is defined as establishing a significant relationship between users and products.

Generating communication between user and product is aimed at interaction design. This communication is based not only on physical features as it is also emotional in its nature. The user is impressed by the successful interaction design in terms of emotional and personal. For instance, painting can be defined as a challenging job. Besides, it is an interactive product in both emotional and personal manner. Moreover, instead of replacing interaction design to technical field or an artistic venture, Human side of technology is emphasized by Interaction design. It is stated by Richard Buchanan, former head of the School of Design at Carnegie Mellon University, as “humanizing technology” (Kolko, 2010):

“Design is not a trivial aspect of the development of information technologies; it is the central discipline for humanizing all technologies, turning them to human purpose and enjoyment.”

Models are widely used in the design, science, and development. Implemented, mental, and represented models are appeared and discussed while designing and developing a product and service. Implemented model is described as a representation that explains how the mechanism works. For example, to stop your car, you have to hit the brake pedal. When you hit the brake, the valve located in the brake booster moves, then the negative pressure inside the intake manifold generates enough pressure to move hydraulic fluid to move brake pistons that are located on wheels. Then the brake piston uses this pressure to generate enough pressure on brake disks with brake pads. Eventually, the car stopped because of this friction. This is called an implemented model, nevertheless, this is not known by most car drivers. Besides, the driver does not need to know how the mechanism functions in order to stop the car. The model in the user’s mind is by far simple. Most of the users think that when they hit the brake pedal, the wheels would stop turning due to the pressure on the brake pedal. This conceptualization is named as the mental model. Every software has a visual interface that is designed by a designer. What is going on inside is demonstrated by this interface, regardless of what is actually happening inside. This is called the represented model. Under these circumstances, the products where the represented model which is closest to the mental model in the user’s mind are easier to use and therefore more desirable. Similarly, this issue is grounded in human-to-SMA interaction. Most users are not aware of how SMA works technically. That’s why the SMAs’ represented

models that are the closest to the user's mental models offer more positive user experience. Besides, it would be successful in terms of interaction design. SMA is software. However, it is accepted as a digital agent by most of the users. Interaction and user satisfaction are affected by the represented model of SMAs (Cooper, Reimann, and Cronin 2007).

User research is interpreted as different things in different disciplines. This user can be a customer, call-center operator, cell-phone owner, and a person who tries to achieve some goal. This goal is done with a software or machine. Moreover, This definition is offered by Robert Schumacher (Schumacher, 2010):

“User research is the systematic study of the goals, needs and capabilities of users so as to specify design, construction, or improvement of tools to benefit how users work and live. “

Additionally, it is indicated famously by Edward Tuffe (Sauro, and Lewis, 2016):

“Only two industries refer to their customers as ‘users’: computer design and drug dealing.”

SMA design is a variety of software design. In order to establish a successful SMA interaction, SMA design has to be based on specific studies that focus on user benefits, requirements, and capabilities. In addition, whether research is conducted qualitative or quantitative, for successful user research, the represented group that is required for intended statements has to be specified (Sauro, and Lewis, 2016). Multimedia notifications of SMA and personification of them by users are evaluated in this study. Beyond the numbers, feelings, emotions, and experiences that are unveiled during human-to-SMA interaction was important for gathering data for this study. Consequently, in-depth interviews were required. That’s why qualitative research was conducted. Moreover, because of the importance of experience, quite experienced users were preferred. Interview questions were based on classical questions such as age, gender, SMA usage time to occurring interesting experiences during interactions, to feelings when SMA acts like a human, and expression of the user when human-like features occurred.

CHAPTER 3: PRELIMINARY FINDINGS OF THE USER RESEARCH

Despite this gender dominance of male participants over woman participants, a more talkative SMA user profile that gives more detailed information about their experiences on SMAs is depicted by the female users.

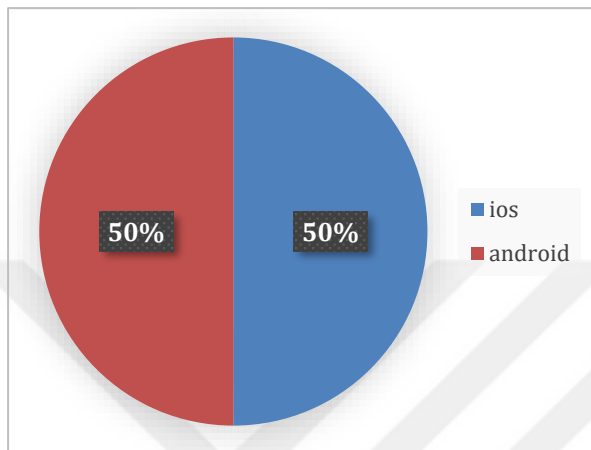


Figure 1. Preferred mobile operation system by user

Whereas Siri is preferred as SMA by five participants, Google Assistant is five equally (figure 1). Default SMA applications were preferred by all participants. None of them had been changed by the user to another SMA application. Even though it is not possible for Android users, both Assistants software can be installed as SMA on iOS.

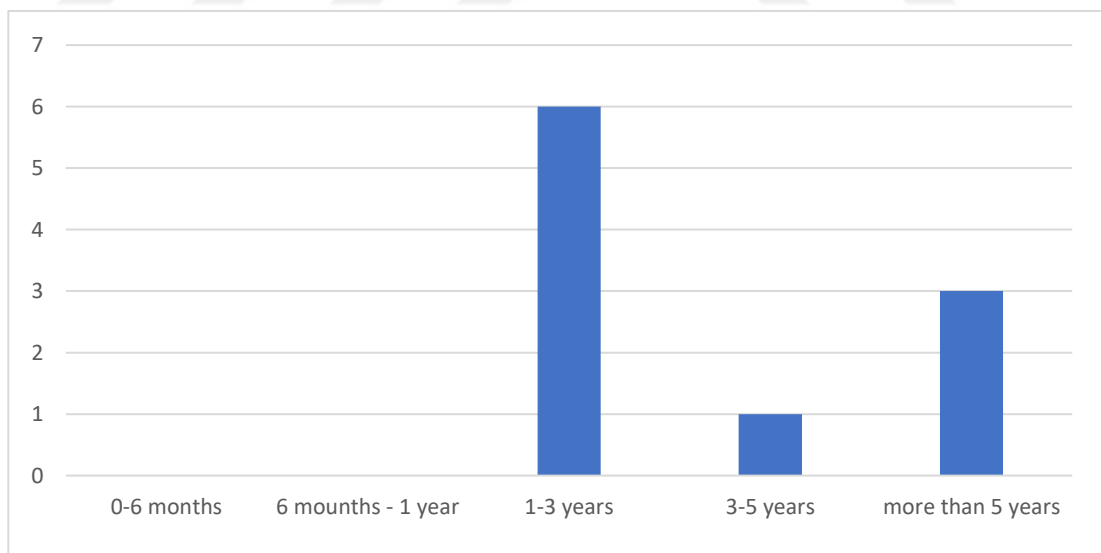
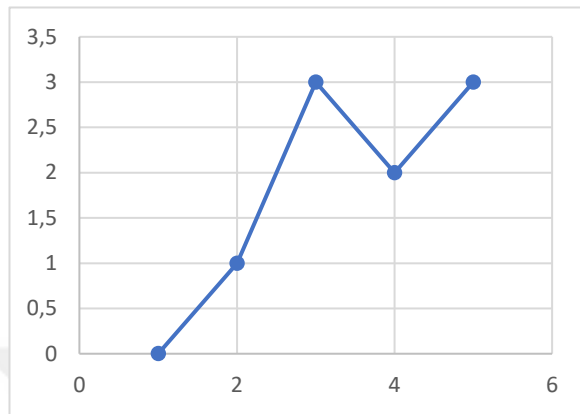


Figure 2. Average experience of users on smart mobile assistant

Most of the participants are late adopters. While six of the participants have reported that their average SMA usage time was between one to three years, three of the participants stated that they have been using SMAs for more than five years. None of

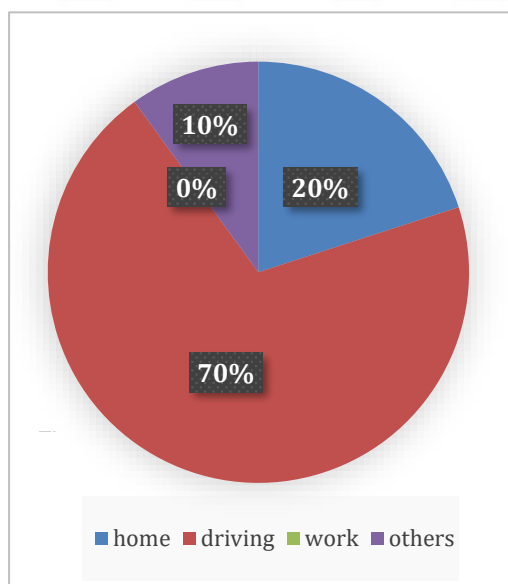
the participants reported less than six months and six months to one-year SMA usage time. (figure 2).

The satisfaction result of participants was calculated in ordinal data value. Satisfaction of SMA was rated from zero to five by the participants. While the lowest satisfaction



was indicated as zero, the highest one was represented as five. Average of satisfaction rate SMA was above three. It was overt that most of the participants were satisfied by SMAs. The lower value was rate two that was given by a participant who was a Google Assistant user, 33 years old, engineer, and female. The higher rates were given by three participants, one of them Google Assistant two of them were Siri users who were at different ages and occupied different jobs (figure 3).

Figure 3. Satisfaction rate of SMA users. Zero is the lowest, five is the highest



SMAs are preferred by participants mostly while driving. Whereas one participant uses SMA while modifying her caravan or carrying bags, two of the participants prefer using SMA at home, SMA was preferred by seven participants in order to assist themselves in driving. The interviews indicated that SMAs were preferred by the users when their hands occupied other jobs and they were unable to touch the device (figure 4).

Figure 4. Location preference where SMA is accessed by user for assist

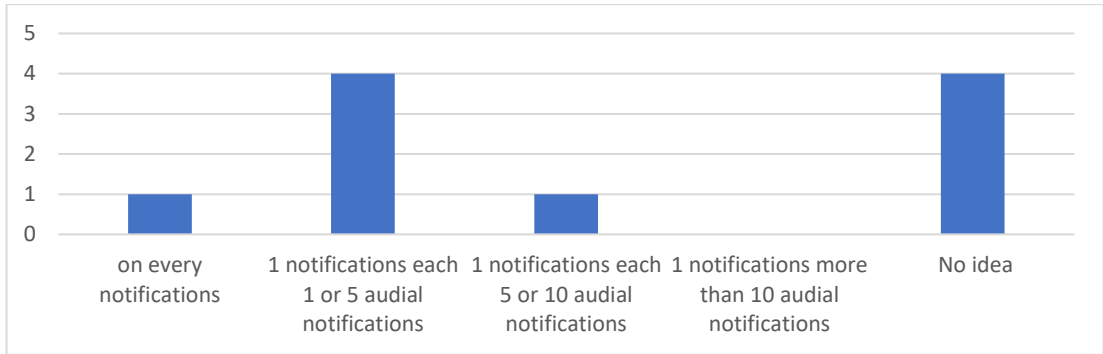


Figure 5. Frequency of visual notifications versus auditory notifications

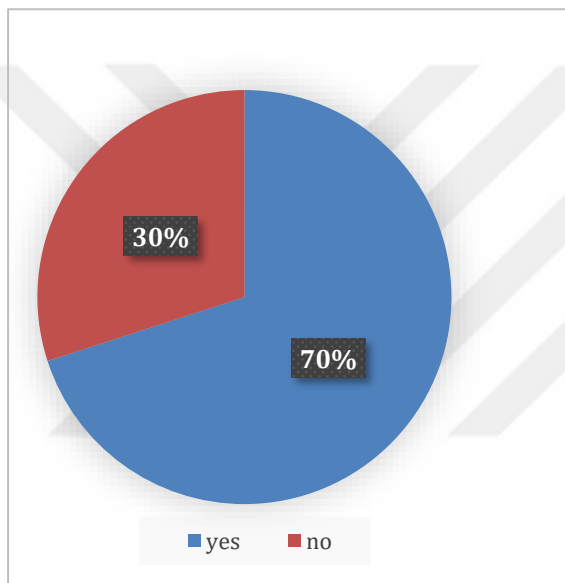


Figure 6. Text and auditory based notifications that is generally accepted as sufficient by SMA users

Visual notifications were considered by some of the participants. As a confirmation, visual notifications were evaluated positively by participants. The existence of visual notification was evaluated positively in terms of offering more clear and vivid interaction. However, it was evaluated as secondary visual confirmation. Primary notification was perceived as a voice by most of the participants. Besides, visual notifications were not considered important by some participants. Besides, absence of visual notifications is not considered to be a vital factor by the participants (figure

6). Visual notifications of SMA were barely detected by some participants, especially those who prefer them in driving. While visual notifications were found necessary by three participants, they were not necessary by seven participants in this research. Moreover, the rate of the importance of visual notifications declined dramatically among participants who prefer to use SMA in driving. However, during the use of navigation software, the perceiving the significance of visual notifications increased.

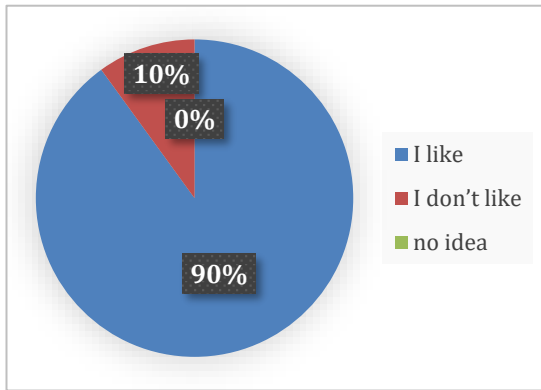


Figure 7. Gratification rate of SMA voice by user

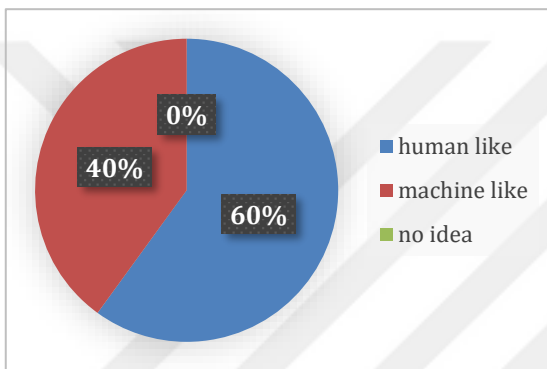


Figure 8. Evaluation of SMA voice by user as human-like or machine-like

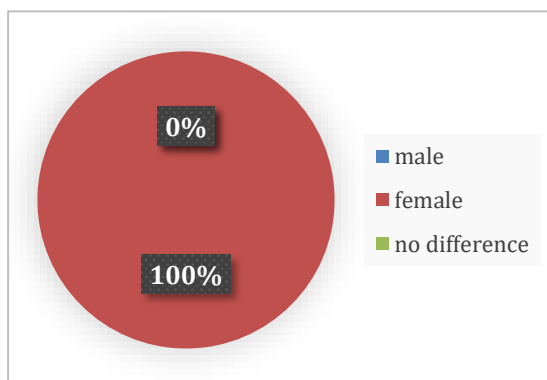


Figure 9. Gender of SMA voice, which is preferred by SMA users

Various questions on the Voice characteristic of SMA were also a part of the interview. The Voice of SMA was liked personally by nine of the participants. While the voice of SMA was reported as human-like by six of the participants, it was found as machine-like by four of them. Siri's voice was appreciated by all Siri users. Nevertheless, Voice of Google Assistant was liked by four participants and disliked by one participant (figure 7). Additionally, the voice of SMA was evaluated as slightly human-like by six of the participants. It was found machine-like by four participants (figure 8). Moreover, both Siri and Google Assistant were rated as having a human-like voice by three participants against the negative vote of two participants.

The female voice was preferred by all participants. It was the default choice determined by the developer. Any complaining or attempting alteration to male voice was not reported by participants (figure 9).

On multimedia interaction, emoji usage by Google Assistant was observed. However, any emoji usage is not reported by Siri users (figure 10).

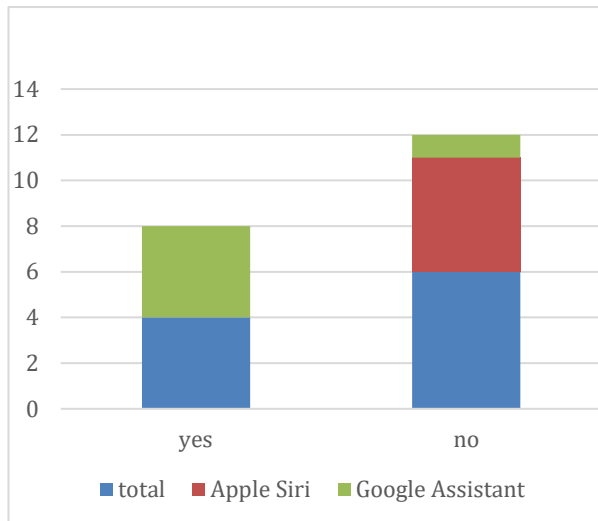


Figure 10. Emoji usage tendency of SMA which is reported by user

one participant. Besides, there was a user opinion that emoji usage does not generate any difference in a user's mind (figure 10).

Emoji is a word that is grounded in two Japanese words. The *e* is stated “picture” and *moji* means “letter”. The study showed that text was not found sufficient by some participants. Therefore, emojis are used for expressing non-verbal implications that were carried by text deficiently (Tang, and Hew, 2019). Consequently, this issue leads to emojis being adopted. It is indicated by a participant:

“It didn’t look pretty when there was only text on the screen.” P11 added that the text looked “weak and powerless, whereas emoji are so much more lively.” P4

Additionally, emoji or stickers can alter the meaning of the text. Emoji usage for altering text is described by a participant:

“Sometimes when it’s not appropriate to say something or when you don’t have a proper word in mind, stickers will help you to communicate. For example, Zhijiang people use ‘the brain-chopper’ to curse people. But if you really use this word, it sounds pretty serious. Use a knife [emoji], then it becomes entertaining.” P1

The tendency of emoji usage of SMA was reported by four of ten participants. While any emoji usage was not stated by all Siri users, existing emoji usage was indicated by four Google Assistant users. However, it was not evaluated as sufficient and appropriate by two of four participants. While insufficient emoji usage of Google Assistant was reported by one participant, there was not an inclusion to emoji usage of Google Assistant was reported by

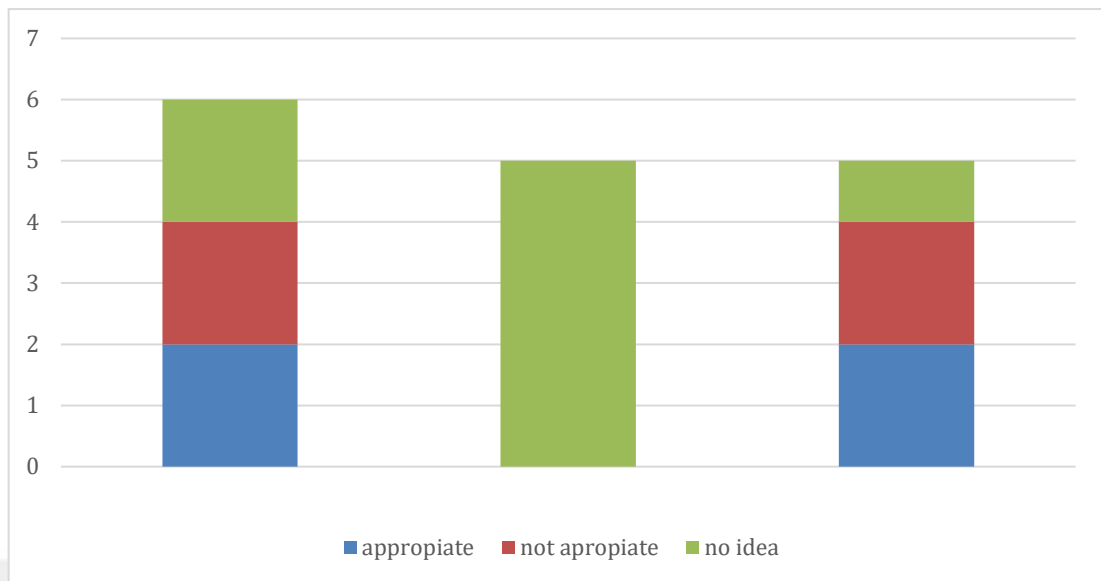


Figure 11. According to user opinion, congruity of SMA emoji usage

The study demonstrates that participants preferred stickers as well as a part of their personal identity. For instance, in another research about emojis a participant whose occupation was police showed multiple saved stickers that related to police and indicated that these were collected from his colleagues. Moreover, P6 stated that the only thought of his emoji usage is a depiction of his feelings precisely. Apart from these, to be sure that the texts were correctly comprehended by the receiver is a motivation factor for emoji usage. It becomes essential especially making conversions at complex and nuanced languages such as Chinese, and, Turkish (Zhou, Hentschel, and Kumar, 2017). In computer-mediated interaction, emojis and stickers play an important role. The reasons behind this are expressing emotions, avoiding misunderstandings and replacing textual statements for enjoyment and to be funny, and for social purposes. The impacts of emoji usage are increasing perceived intimacy in communication. Friendly emojis tend to be perceived as more favourable, expressing the intended meaning of the message, and mitigating negatively or fostering positivity.

According to interviewed 30 participants about their emoji usage, the motivations are fostering personification, expressiveness, humour, and amusement (YING TANG, 2019). Besides, some studies demonstrate that using emojis and stickers could help precisely understand the ironic meaning in a message (Thomson, and Filik 2016).

Even though any emoji usage was not reported by Siri participants, It was faced by many Google Assistant users. Emojis were used by Google Assistants mostly for expressing facial gestures such as smile, anger, or sadness. In contrast, using emojis for fostering the meaning of existing situations by Google Assistants was reported by just one participant. Using emotion expressing emojis was perceived by one participant as chatting with a real person:

“Google Assistant really texts just like us.”

P3 (A Google Assistant user)

According to the user interview, the functionality issue of SMA was mostly a complaint part of SMA experience. Capabilities and doing a task in the right way are discussed as functionality among the participants. The functionality of SMA was not found sufficient by six of ten participants. It was indicated by one Google Assistant user that in case of changing somebody's personal information in contacts, Google assistant has demonstrated a link and said: *“You can change this on the following link.”* instead of changing this information. Moreover, Google Assistant has missed information of a person in contacts. For example, the user saved a person as “husband” in the contact list. In the beginning, when the user gave an order to Google Assistant, “call husband” the husband was called by Google Assistant. However, after a while, this order was not comprehended by Google Assistant. Then the order has to be changed to “call darling”. This issue was profoundly criticized by the Interviewed user. Apart from that, misunderstanding of orders was indicated by many participants. Calling the wrong person or dictating the wrong words were indicated by some of the participants. In this case, displaying comprehended sentences by SMA that is dictated by participants and demonstrating visual clues of what is proceeding is not only



Figure 12. During interaction, emojis are frequently used by Google Assistant

essential but also life-saving features. Visual confirmations were evaluated as essential by one participant because of understanding and interfering with unwanted actions. For instance, the picture of a called person can be displayed visually by both Siri and Google Assistant. In case the wrong person is called, the user can be aware of this issue and intervene in it.

There were subtle differences between Google Assistant and Apple Siri about issues that were related to calling someone that was reported by the interviewed participants. However, any issues with calling the wrong person were reported by Siri users. Additionally, hints and clues about usage were advised by one Siri user. In contrast, the hints that can be ordered to Google Assistant were frequently displayed on Google Assistant application interface (figure 13).

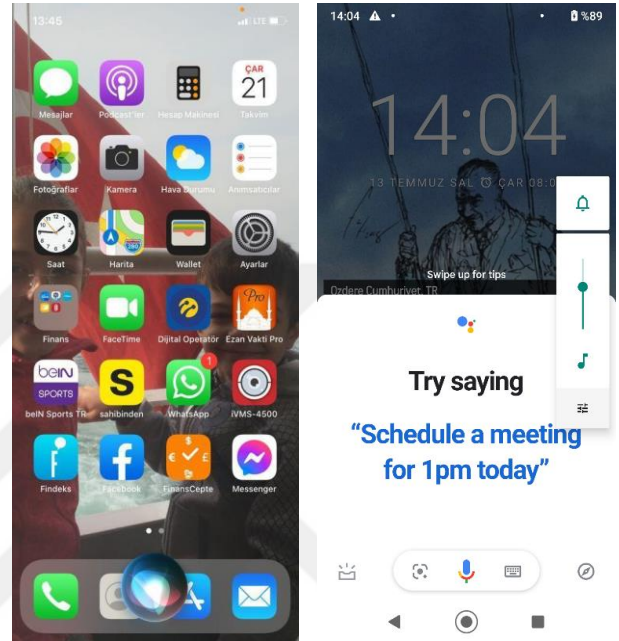


Figure 13. Hints are displayed on Google Assistant interface when it is launched (right picture). The circle is appeared on the screen, when Siri is activated (left picture).

SMA was often preferred during hands-free interaction. These were widely used by a user in a car.

At a computer expo (COMDEX), Bill Gates reported comparing computer industry with the automotive industry:

“If GM had kept up with the technology like the computer industry has, we would all be driving \$25.00 cars that got 1,000 miles to the gallon.”

it was responded by General Motors:

“If GM had developed technology like Microsoft,The airbag would say ‘Are you sure?’ before going off (Mikkelson, 2010).

In contrast to the automotive industry, confirmation is important in human-to-computer interaction. Users are faced with many confirmations during computer usage. When saving or closing a document, shutting down the computer, and writing something on a text editor, confirmations are always visible. Google search engine is mostly known with confirmation *“Did you mean”*. The user is informed by confirmations via various multimedia notifications. They are seen as pop-up on the interface visually, repetition of action as auditory, and vibrations on hand-held devices.

Working principle SMA is based on text dictation. As it is mentioned in the “introduction” the voice of the user dictates to texts and text-based information is sent to a specialized server. During voice-to-text conversion, some similar words can be chosen, and the wrong order can be sent to the server. Besides, because of some software bugs, the order can be interpreted differently from the user’s intended command, unwanted results can occur. In this case, the user can comprehend this wrong proceeding, and interfere.

Moreover, confirmation of SMA leads to be personified. If a human assistant is imagined, she/he always gives confirmation. When the order is received, he/she smiles or shows a facial expression that it is understood, and the required task is confirmed by repeating with his or her voice.

The most visual notification was reported by the participant was textual confirmation. The auditory response is repeated by text-based notification. The user experience about it was mostly positive. Text-based notifications are generally used for confirmation by SMAs. The misunderstood command is noticed by the users with text-based confirmations. Reducing misunderstanding by the contribution of text-based notification was reported by many participants. Moreover, demonstrating celebrity faces when it was asked, currencies with a visual number value, showing search results

when it was used for web search, were some of the visual notifications that were indicated by participants.

Even though any emoji usage was not reported by Siri participants, It was faced by many Google Assistant users. Emojis were used by Google Assistants mostly for expressing facial gestures such as smile, anger, or sadness. In contrast, using emojis for fostering the meaning of existing situations by Google Assistants was reported by one participant. Using emotion expressing emojis was perceived by one participant as chatting with a real person:

“Google Assistant really texts just like us.”

P3 (a Google Assistant user)

According to user interviews, Google Assistant was personified by users more than Siri. Sadness was expressed by one Google Assistant user when insulting Google Assistant:

“I was sad, because, she was laying a guilt trip on me.”

P4 (a Google Assistant user)

Happiness was reported by one Google Assistant user in case of a compliment between user and SMA:

“We are complementing each of us. I have a good relationship with her.”

P3 (a Google Assistant user)

Similarly, Siri is faced with similar things which come from its users. However, adoration of technology is the most indicated opinion among Siri users. As an exception, marriage is proposed by one Siri user. Before this proposal, the participant quarreled with his girlfriend. Any conversion was conveyed between this couple. Participants preferred making conversation with Siri. It was indicated by the user:

“I wasn’t used to making conversation with my girlfriend. So I thought, let’s talk with Siri.” P7 (a Siri user).

CHAPTER 4: EVALUATING USER EXPERIENCE ON SMA INTERACTION

4.1. Demonstration of Differences and Similarities Between the Notifications of Apple Siri and Google Assistant

The user interacts with the device by helping with notifications and feedback. These are fundamental elements for establishing a successful interaction design. When icons or links are clicked they are glowing or generating sound effects. These feedbacks are essential because the user needs to understand what is going on and how they are doing. Similarly, some notifications and feedbacks are generated by SMA's. The notification of SMA is mostly known as voice. it is not only most known but also distinctive. Even though there are some smart assistants which have limited visual notification abilities such as Google Home, Amazon Alexa, and Apple HomePod, Blinking led lights on them are generated as visual feedback. However, SMAs have a chance to give sophisticated visual notifications to the user. In this chapter, the multimedia notification varieties of Google Assistant and Apple Siri and similarities and differences among them are evaluated. In the course of the research, User interviews are conducted. The experiences, memories, complaints, and appreciations are recorded. Finally, these data are combined with a literature review in order to get more detailed data.

4.1.1. Evaluating Variety of Multimedia Feedback of SMA During Interactions to User

Multimedia notifications are abundant in SMA's. The most common feedback of SMA is the voice which is followed by visual feedback. Even though both visual and auditory notifications are used by SMA, slight differences can be observed. In this chapter, these differences and similarities are evaluated. This study is grounded in user interviews and literature reviews.

During the interaction with SMA, multiple ways of feedback and confirmation are vital because of two reasons. Firstly, the user is mostly occupied with other jobs or tasks. That's why he/she may not fully focus on SMA's process. Secondly, commands can be misunderstood and applied as the wrong command by SMAs. Therefore, text-based confirmation is evaluated as essential by most participants. Consequently, attempting

complicated tasks is avoided by the user in order to face high social costs or embarrassment such as calling or sending a message to someone.

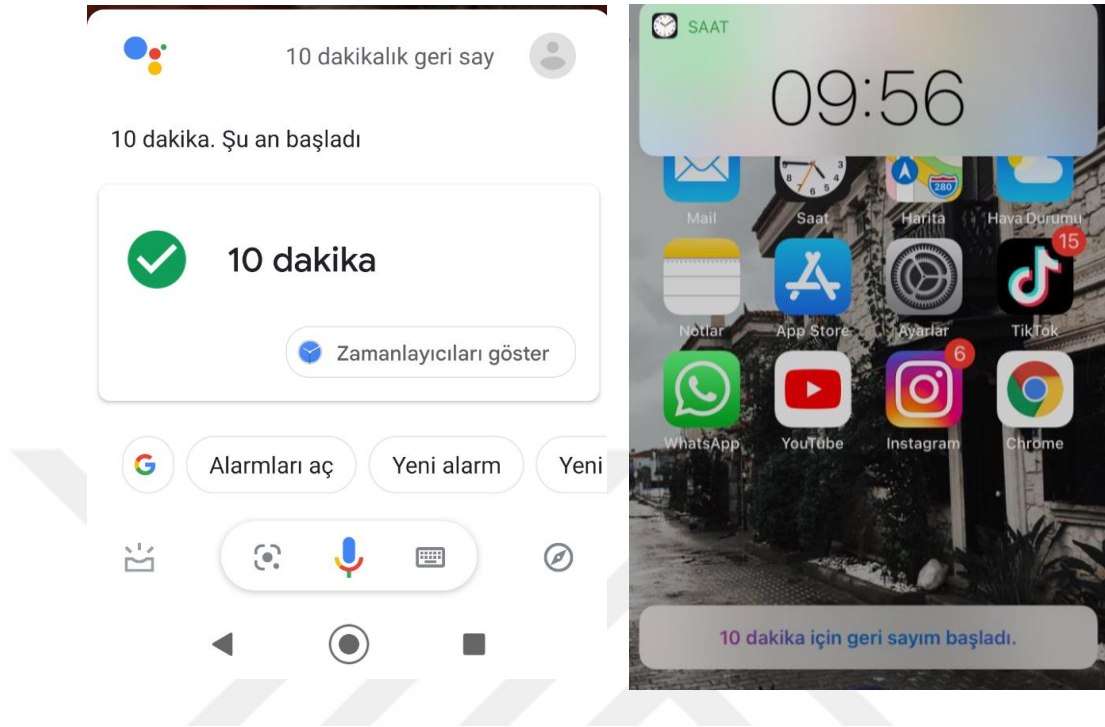


Figure 14. Provided visual and auditory confirmation notifications by Google Assistant and Siri

As a result, notifications play an essential role in ensuring a positive user experience. Communicating with multiple senses makes them easier to understand. Human to human conversation is not only based on speech but also facial expressions and gestures generated by people that are monitored by a listener. At that point, many emojis that express emotion and foster the meaning of sentences are preferred by Google Assistant. The interface of Google Assistant resembles a chatbox according to many Google Assistant users. Inherently, this design choice helps Google Assistant for using emojis and facial expressions. However, the same design choices are not obtained by Siri. Additionally, there are some user recommendations about visual, audio, and tactile feedback of SMA during the interaction session. Indicated by participant:

“I asked Alexa a question, realized she was unresponsive [...] It should have a resting light or I am alive.” A participant (Lopatovska et al., 2019).

Device feedback is essential in successful interaction design. Jakob Nielsen mentions this fundamental design principle in his book called “10 Heuristics for User Interface Design”:

“The system should always keep users informed about what is going on, through appropriate feedback within a reasonable time” (Nielsen, 1994).

In this case, Apple Siri and Google Assistant have similarities and differences. First, while listening to the user, progression feedback is given by both Apple Siri and Google Assistant. Whereas the waving bar is displayed by Siri, this feedback is given by Google Assistant in the form of a colorful stick. The colours that are used in Google Assistant’s symbol are designated from Google's corporate identity and its colour scheme. In these circumstances, it is stated that Google Assistant bonds a connection with its brand via colour. Second, mainly text and voice are preferred by Apple Siri and Google Assistant in order to interact with the user. However, the text interface of Google Assistant resembles a classical chatbox. It is close to a chat box where people use it to chat with friends. Additionally, the combination of visual and textual interaction is evaluated by the users as a more interesting and vivid communication. Misunderstandings and preventing broken communications are reduced by the help of visual aids. It is helpful and important for people, especially those who use SMAs in a language that is different from their native one. In this term, while interacting with the user, the user experience is enhanced by Google Assistant via emoji usage. Another difference is the visual interface style. Dark text on a white background is preferred by Google Assistant, while Siri uses its opposite. Speed and jump length are two essential factors in terms of characteristics of dynamic information for the leading-display design. The speed of the leading display is calculated by the number of words that are displayed on the screen per minute. The measurement unit is words per minute that are abbreviated as “wpm”. The amount of characters displayed on the screen is measured by jump length. jump length is the length of characters that are displayed step by step on screen. For example, When a new character is unveiled, if there is a second character that can be generated in subsequent seconds for a composing word then the jump length is the length of one character.



Figure 15. Example of two leading displays with the same speed but different jump length settings (Source: Wang, Fang, and Chen, 2003)

Under the same speed setting, smoother movement is performed by smaller jump lengths. In figure 15, two displays with same speed leading but different jump length are demonstrated (Wang, Fang, and Chen, 2003).

Text is presented on mobile devices in various formats such as paging, scrolling, leading and rapid serial visual presentation (RSVP). In the smart mobile assistant experience, RSVP scrolling, and rarely leading are preferred methods for demonstrating texts. However, there are some differences among them in terms of ease of reading, comprehensibility, and taking effort. According to the studies, the most efficient method for reading speed was paging. 217,7 wpm Reading speed was reported for paging. It is followed by leading with 195,2 wpm, scrolling with 178.1 wpm, and RSVP with 135.4 wpm. When comprehensibility is evaluated, the efficient methods were RSVP and scrolling with a 92.5% intelligibility ratio. These are followed by leading with 88.8 %, and paging with 87,5%. Additionally, according to nasa-tlx scores of all four methods, the least mentally demanded method was paging. It is followed by RSVP, scrolling, and leading. Moreover, conforming to eye movement analysis, the least eye movement was reported in RSVP format. In conclusion, the faster method is paging according to these results. However, the most comprehensible ones are RSVP and scrolling. In SMA experience, current dynamic text presented methods offer the slowest but more comprehensible and less effort-

required design solutions. Besides, it was asserted that readability is increased by RSVP (Öquist, and Lundin, 2007).

In addition, a study (Bernard, Chaparro, and Russell, 2001) as a result of comparing word-by-word, three-line, and 10-line text presentation methods, word-by-word and 10-line format was evaluated by participants as a method that provides better reading than others. Moreover, sufficient information of context is not provided Smaller snippets of text. That’s why, more cognitive pressure for comprehending text might be felt by readers. Consequently, sentences and paragraphs are indicated as preferred formats by participants in indicated questionnaires (Sharmin, Špakov, and Rähä, 2012).

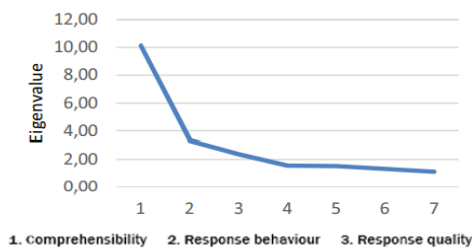


Figure 16. Functional analysis of user evaluation data (Source: Klein et al., 2020)

The most comprehensible method is dark text on white background with 250-300 wpm text speed. However, jump length is also important for readability. In this case, the most readable jump length is between 0.35 and 0.7cm (Wang, Fang, and Chen, 2003). Nevertheless, higher jump length text is observable in both Siri and Google Assistant interfaces.

Instead of showing words letter by letter, the whole word is displayed on a screen, and then sentences are formed. The study shows that the highest factor value when evaluating SMA is comprehensibility. The highest-ranked factor is comprehensibility is shown in figure 16. Even many SMA’s are evaluated and rated in this study (35 Alexa, 27 Siri, 26 Google Assistant, and 8 Other) this depicts the overall evaluation of SMA (Klein, 2020).

4.1.2. Language Support

Even though the variety of language input available on both Siri and Google Assistant, many more languages are supported by Google. The number of supported languages by Google Assistant are more abundant than Siri. While over 30 languages are supported by Google Assistant, Siri does 21 (Wiggers, 2019). Even though people’s

native language is not understood and responded by SMAs, it would be a more frustrating experience when these are used by foreign speakers with different accents and utterances. However, in this research two Turkish participants stated that they prefer the English version of Google Assistant due to ease of communication. However, Siri's accent and cultural adaptation were also appreciated by a few participants. Siri was admired for both pronunciations of accent and having cultural knowledge. A participant stated:

“I switched her over to Japanese and she is a lot more serious and polite, and less likely to crack jokes. I think there is a personality for each of the different languages depending on what that language requires I guess (...) unlike the English version, the Japanese do not really do that, it is not colloquial in the same way, it is very formal. It's very much an assistant rather than a personality on your phone if that makes sense.” (Cowan, 2017).

Beyond language support, answers of Google Assistant are evaluated more accurately than other SMAs. Voice variants are more easily comprehended by Google Assistants. The dominance of Google Assistant existed in voice recognition and human free interaction with a 59.8 % achievement rate. However, Siri responded with 43.98% accuracy. Consequently, even though there were subtle differences between them, responses of Google Assistant were more accurate. Moreover, not all daily-based questions were responded by SMA. Responded questions were composed of 17.35% daily-based questions. While 45% of these questions were responded to by Siri. 59.80% responded by Google Assistant. It was obvious that Google Assistant was more successfully engineered than Siri (Tulshan, and Dhage 2019).

4.1.3. Analysis of User Reaction to Multimedia Notification of SMA

User to software interaction is based on the relationship between notifications or feedback of software and given inputs by the user. Demonstrating reactions of users is an essential part of this interaction. Notifications and feedback that are provided by SMA are reacted by the user. Users may be satisfied with results, give another order, and express emotional reactions such as anger, happiness, etc... Valuable information about user experience is carried by these reactions. That's why user reactions to SMA

notifications and their memories are collected and evaluated via interviews for this study.

In both Google Assistant and Siri, the user receives interaction feedback via various multimedia notifications. They are mostly known as speaking digital agents. Consequently, primary feedback is voice. This is the most distinctive feature that assists SMA to be personified. Most of the participants prefer to hear low-status female voices. In the study, while high-status languages were indicated to passive and formal sentences, low-status language was composed of active and informal sentences. This is observable in both SMA users. In general, multimedia notifications type and frequency were indicated as adequate by most participants. Most of them did not complain about it. However, Siri was criticized for not giving enough voice feedback in response to some questions. For example, when asked to deliver the weather forecast, Siri said “here is the weather” and the weather forecast was displayed visually. Nevertheless, sometimes the device cannot be reached by the user to unlock the screen at that moment. This became a frustrating experience. Besides, some of the users complain about this issue:

“Making it more hands-free because if I ask the weather, Siri does not read it to me, it just goes ‘okay, here’s the weather,’ and I still have to walk over to my phone and check.” (Lopatovska et al., 2019).

However in the time frame of this study, interviewed participants who prefer Apple Siri reported that Siri has voice feedback while reporting weather forecasts. It seems that this issue has been fixed after receiving a complaint about Siri. However, the voice volume of Siri was criticized by some of the participants of this research for not being audible.

Siri and Google Assistant have different voice characteristics. Mostly, Siri’s voice was evaluated by the user as having a more human-like voice with the latest updates. Therefore, user experience on the personification side was fostered. Some of the participants indicated that they feel bad when they insult their SMA due to a mistake because afterwards, the SMA apologizes.

Similarly, an Amazon Alexa user said:

“Alexa makes it sound like a person, so when it makes a mistake, I feel mad and I yell at it. I feel like I’m being rude to a child. When I changed its name from ‘Alexa’ to ‘Computer,’ I’m not mad at it when it makes mistakes.”

Murtaza, Participant (AnswerLab, 2018)

A study shows that Google Assistant responses were found slightly better by participants while considering being positive of answers. The researcher linked this result to the voice of Google Assistant. It is female voiced and has a tendency to express behaviour that is more natural and more emotional than others do (Berdasco et al., 2019).

The second common feedback of SMA is visual notifications. It is composed of text-based communication or web search results with visual content such as images, videos, or interactive maps. In this term, some SMAs such as Google Assistant is seen as chatbot due to texting like a real person. Moreover, during interaction emojis are used by Google Assistant frequently in order to express feelings and improve statements. According to user interviews. These were not much different from humans in terms of emoji usage. To put it another way, it resembles that human-to-human chat conversation.

Apart from these, vibration is used by Google Assistant as feedback. Even though it is rarely used. In case the device is locked or Google Assistant is not launched and if Google Assistant is intended to be launched as called (“Hey Google”) or pressing the microphone icon on the Google Assistant interface, the device vibrates in order to report to the user that Google Assistant has activated.

4.2. A Study on the Possibilities of Improving User Experience through SMAs’ Multimedia Feedback

Many data have been collected from both literature reviews and user interviews. This data does not only include user satisfaction, but also how negative experiences occurred and were fixed. Despite the vast improvement of SMAs in the past decade, it is still defined as primitive artificial intelligence. More intelligent and dangerous AI

is depicted in science fiction. It is in the near future or not, it is inevitable. Considering these potentials, a vast improving chance of user experience would be possible. In this chapter, which improvements might be applied and may be developed are discussed.

When a new technology has been released, some people are willing to own this technology due to its potential. These people are called early adopters. New technology and functionality are the primary factor to buy it for early adopters. Within this period, marketing research is done for detecting required developments via customer surveys. Then, it is reported to the company to be realized. As a result, new products are announced with new features and technologies.

When the fundamental features of this technology are provided by all of the manufacturers around the world, this technology will be sufficient to provide the fundamental needs of the user. Besides, it becomes more stable and reliable. For instance, going somewhere is provided by a car regardless of its quality, equipment and performance. Beyond their basic features, the other features are started to be evaluated by the customer such as price, appearance, prestige value, and maintenance. At that stage, technology starts to mature. When this technology matures, customer profile and user needs change. Also, the brand starts to be considered as well as the actual products and its features. Additionally, it is observed that the names of the companies which offer this technology begin to change. The technology indicative words in the name of the company are removed or changed. For instance, The “Apple Computer Company” and “Apple Computer Inc.” were used to be used as company names by Apple until 2007. The name of the Company has been changed to “Apple Inc.” in 2007. When the technology is mature, a new customer profile is unveiled. These are called “late adopters”. Whereas the early adopters want technological development and ignore cost, late adopters require reliability and simplicity. Late adopters are not like early adopters. Their tendency on using a product is “turn it on, use it, and forget. it.” Their world is completely different from the early adopter’s world. Additionally, the requirements of the consumer world are different from the high-technology addicted world. Finally, the technology reaches adulthood. Cost, appearance, reliability becomes more important than its technological features. For example, a wristwatch can have less function or be less accurate. However, it can be sold at a higher price than a more accurate and featured one. Because emotion is sold

more than technology in matured technology. Unfortunately, the same thing can not be indicated for computer technology.

If a company wants to grow and enter a new market, it has to provide late adopters' needs and their requirements. The product, service and technology have to be transformed from technology-oriented to customer-oriented. Instead of technology, users have to be aimed in order to improve products, services, or technology. The technology is designed as human-centred. Convenience and reliability are considered more important than technological superiority. In other words, it has to be developed by the principle of Human-centered design principles in order to be successful in the market (Norman, 1999).

SMA's are not close to maturity. Users who are early adopters, and technology addicts have more positive experiences on SMA interaction than users who are late adopters. Even though there are some attempts, It demonstrates that SMA technology still stuck on the technology-oriented side, not human-oriented.

Showing the right information at the right time is a basic feature of a smart assistant. This would provide a positive user experience. This issue is handled more positively by SMA today than in the past. Nevertheless, improving user experience by increasing usability, utility, and interaction aesthetics are not enough today. In order to create differences and offer more positive experiences and services that personify user's offers and develop a system that fosters understanding of user's needs and the world is started to be focused by developers and designers (Dove et al., 2017).

The primary motivating factor for using SMA is saving time. Giving commands by talking is by far faster than opening applications, tapping and typing. Generally, users are happy with hands-free usage ability. A participant stated that:

“It's brilliant that it reads messages to you and gives you the option would you like to reply...automatically, without pressing the button again it goes into what would you like to reply and it's almost like you're having a conversation with your co-driver”

Andy

However, in order to launch SMA, devices have to be unlocked by the user. Today, generally typing pin code or showing fingerprints to devices are preferred by users to unlock to devices. When the hands of the user are busy, leaving what jobs users are doing and launching the SMA is mandatory. In this circumstance, SMA usage is given up by the user and thought that it is useless and rubbish:

“If I have to unlock my phone to use it (Google Assistant) with my finger, why should I use SMA.”

P1 (A Google Assistant user)

As long as the hands-free interaction can be activated seamlessly, the users are mostly happy with their SMAs. It is overt that one of the most preferable advantages dictated by participants is hands-free usage ability. SMAs are preferred by many users while their hands are occupied by other jobs such as driving, child caring, and carrying something. A participant stated that:

“I think that it is very useful while I am carrying shopping bags.”

P4 (A Google Assistant user)

Additionally, the time saving aspects of using SMAs is another motivation factor. Not only for daily jobs, but also the business aspect is evaluated as important by many participants. Spending time for a certain task is reduced. A participant stated:

“My feeling originally with Siri was that it was a toy....you’d ask it to do stupid stuff and then you start to do certain things with it and it starts to work, you know, like putting stuff in your calendar, and then it just becomes like an easier way of doing things.”

Mike (Luger, and Sellen, 2016)

“I’m also constantly asking Siri to set appointments, reminders, alerts and alarms because on the iPhone each of those takes, you know, 4 to 7 steps, not including typing. So Siri is good for those things that would otherwise cause me to go to the keyboard.”

Human-to-human communication is not composed of just words and sounds. Humans follow facial expressions and gestures to each other during conversation. This is crucial for communication, which is made more comprehensible. However, SMAs suffer a lack of reading and demonstrating facial expressions and gestures. Reducing multimode interaction to voice-to-voice interaction causes not only unnatural interaction but also difficulty. Besides, Users can comprehend SMAs from the emojis that are used by SMA and visuals in their interfaces, but SMA cannot. However, the expected abilities of a human assistant is to understand user needs, feelings and to foresee user expectations. Therefore, it is against the mental model of the SMA in the user's mind, and even though it is just a software, the user tends to personify and add personality to it. Frequently, users are not programmers. Besides, understanding how SMA works is neither expected nor required from the user. Their tendencies on SMA is replacing them with a human personal assistant. A participant from interviews stated:

“Once I insulted Google Assistant because of its mistake. Then, she apologized in a naïve way. Then I felt upset. I was remorseful. “

P4 (A Google Assistant user)

Conversations are mostly turn-based. Humans wait for their turns to speak. Adding “wake words” to the beginning of every command is found weird and time-consuming by many users. Even though follow-up mode is used by Alexa to handle this problem, but it is not a perfect solution. This issue is struggled by many users and an unpleasant experience is unveiled because of it. One participant stated:

“What I don't like is that [Alexa] doesn't shut up when I start talking to her. This is what more human interaction should be. [...] It would be ideal if it interacted to something less than 'Alexa, stop' — something like 'ok', or 'enough', or pretty much anything that I mutter [...] It's like talking to someone who just goes on and on, and you're waiting to find a pause so you can somehow stop them.” (Budiu, and Laubheimer, 2018).

In the future, user voice tone, gestures and facial expression would be monitored by SMAs for better conversation and subsequently generating more positive user experience (AnswerLab, 2018). This issue is discussed by the cinema industry too. For example, personified smart objects are discussed by a Netflix series “Black Mirror” in season five, episode three. An adolescent young girl (Racheal) has a smart doll that is named “Ashley too”. At the same time, this smart doll has a visual similarity and it is inspired by the singer's behaviour. The relationship between the doll and Racheal is mentioned during the episode. The doll has a personality that is added by the manufacturer. Beyond this, Racheal establishes a friendship between the doll and herself. Moreover, the personification of the smart doll by Racheal and her sister is considerable for predicting foreseen future expectations of SMA. Additionally, it is not only just related to artificial intelligence that the doll has, but also it is related to brand personification. The singer Ashley too is a brand herself. Therefore, it eases the personification of the doll Ashley too.

In human-to-human communication, nonverbal communication plays an important role, as well as verbal communication. Nonverbal communication is described briefly as including all behaviours that are not composed of words. In addition, nonverbal communication is conceptualized as a subset of non verbal behaviour. Communication is described as “ a dynamic and ongoing process whereby senders and receivers exchange messages” by Burgoon, Buller, and Woodcall (Guerrero, and Floyd, 2008). Besides, the potential of nonverbal behaviours are noted by Burgoon and her colleagues. In addition, nonverbal cues such as gestures, touch, and vocal tone are evaluated as high value in communication potential (Guerrero, and Floyd, 2008).

The tone in the human voice helps to understand the meaning and feeling of the user and an essential aspect of language itself. Another participant said:

“I almost feel like a robot when I’m asking questions, because I have to say it in such a clear and concise way, and I have to think of it so clearly. When I try to give a command or ask a specific question, you don’t use much inflection. It’s really just picking up words, it’s not picking up emotions in your voice.” (Budiu, and Laubheimer, 2018).

Hands-free interaction ability to SMA was appreciated by many participants. However, in order to launch SMA devices have to be unlocked by the user both on Siri and Google Assistant side. This issue was complained about by many participants. A participant stated:

“When I was trying to send a message to my brother, it made me unlock my phone, which was annoying because I'm trying to be hands-free, I'm not trying to unlock my phone. And if I was just doing it while not looking at it, (...) if I was in a car or something, I wouldn't be looking at it so I wouldn't know that it hadn't sent the message. And then if I tried to edit the message, it just has me redo the whole thing, which I thought was stupid.” (Cowan et al., 2017).

However, this issue is maybe obviated by new technological developments. The device can be unlocked just by scanning the face of the user with assistance of face recognition technology. Face recognition technology has been developed since the 1960s. With the development of the “Viola-Jones algorithm” the first implementation of face recognition on handheld devices was applied by Samsung with note 7 in 2016. In addition, Apple shows a revolutionary approach in face recognition technology with depth scanning. Apple's depth scanning addition to this technology further advances low light usability and security. Face recognition with depth scanning has been launched with the iPhone x in 2017 (Tilman, 2021). Nevertheless, there are some limitations. When a device is in a pocket or somewhere else the user's face cannot be seen visually. Hands-free SMA experience may be fostered with a helping combination of voice recognition, and face recognition technology.

4.3. Effects of Personification on User Experience

4.3.1. A Study on the Personification Ability of Human Psychology: Heritage or Imprecation

Humans are one of the most social-living creatures on planet earth. Significant traces of the human brain have been left by living in a society. Differences between “us” and “them” are fostered. People, animals, even some inanimate objects which are located in a person's society and family have gained meaning. In this chapter, the human's personification ability and personification side of SMA is evaluated. During this research, user opinions from conducted interviews are evaluated.

SMA is known as software by the users. It is thought by many people that insulting machines and forgiving them imply a mental health problem. However, because of millions of years of evolution. Human beings have evolved into living creatures who are more social and live in communities. Besides, juxtaposing our brain and nearest cousin (Homo Neanderthalensis), demonstrates that the neanderthalensis brain was larger than homo sapiens. Nevertheless, just the backside of the neanderthalensis brain is larger than homo sapiens. The backside of the human brain is responsible for vision and fundamental life skills. However, the social and language skill of the homo sapien brain is larger. This is accepted by science as one of the key factors of our surviving skill despite this our psychical weakness. Homo sapiens are cheated by this life-saving ability of the brain. So, it makes Homo sapiens inclined to personify inanimate objects.

Personification is a common human tendency. A human who names their stuff and talks to them is overt regardless of the human's mental health. Inanimate objects tend to be personified by a person who has normal mental health. Moreover, this tendency of humans is exploited by marketing strategies. Consequently, humans are surrounded by smiling, dancing or talking products. The effect of personification on customer choice is observable in customer experience. Therefore, screaming shoes, own-minded shopping trolley, etc... are used for describing this experience by customers:

"I felt like the sandals were screaming 'buy me, buy me'. The shop assistant could see how delighted I was as she handed the sandals to me.

(female, 20) (Brown, 2011)

Similarly, being in connection with not only the products of a brand but also the brand itself is established by customers. In case a person feels and thinks: "this brand is like me or I am like this brand." eventually the brand embodied a personified character.

This is examined in brand personification. Brand personification is created and presented by a brand with human-like ability. On the other hand, there are many aspects of personification. One of them is brand personification, which is personified by the customer (Cohen, 2013). In this case, because of its name, colour, the relationship of its developer companies, and relationship with other mobile

applications that are already installed on the device, Google assistant is more brand personification oriented. Even though users have to say “ Hey Google” in order to interact with Google Assistant. In contrast, Siri is evaluated as having unique characteristics regardless of its developer company by all of the interviewed participants. Also its name is different from other Apple products such as iPhone, iPad, and iMac even further.

Additionally, there were many complaints about Siri. Siri was criticized by many users due to insisting on using only “Apple applications”. A participant mentioned:

“Because right now Siri's limited to only Apple products, but I hate Bing, and I hate Apple Maps, and the only person that has iMessage is my brother, so it's sort of limited by its Appleness. (...) If she used Google Maps, I might be more likely to ask her for directions if I'm going somewhere but she doesn't, so I don't. (...) and to add Messenger interfunctionality, because I have like, six messaging applications on my phone. If I could say, send a telegram message to Marcus, send a WhatsApp message to my dad, send an iMessage to my brother, because everybody uses different platforms. If I could tell Siri to use a certain application to send a message I would probably use it more”.

However, most of the participants who have taken part in interviews disagreed, as most participants have stated that there was no pressure exerted by Siri on the customers to force them to use developer’s own applications. Technically, Google Assistant is not very different from Siri in terms of compelling users to prefer the services of the developer. However, Google applications are preferred more than Apple applications by both iOS and Android users. Consequently, this issue is faced less frequently by Google Assistant users. While Google Assistant is evaluated as personified with no personification or like an ice-cold-librarian, Siri is very funny and sometimes sarcastic. Even though Google Assistant was ridiculed by some of the participants by imitating its dull voice in this study. The task-oriented character of Google Assistant is reported as more dominant than Siri. On the other hand, even though it is limited and not much more than Siri, Google Assistant has humour ability. For example, when repeating words are requested from Siri, this demand is rejected by Siri by saying "I am not a parrot". Distinctly, this request is accepted by Google Assistant. Also, it is coherent to

its task-oriented characteristic. However, the example of repeated sentences that is given by Google Assistant is satire:

"If you say just behind repeat after me command. I would repeat whatever you say such as repeat after me, I am a moron"(figure 17).

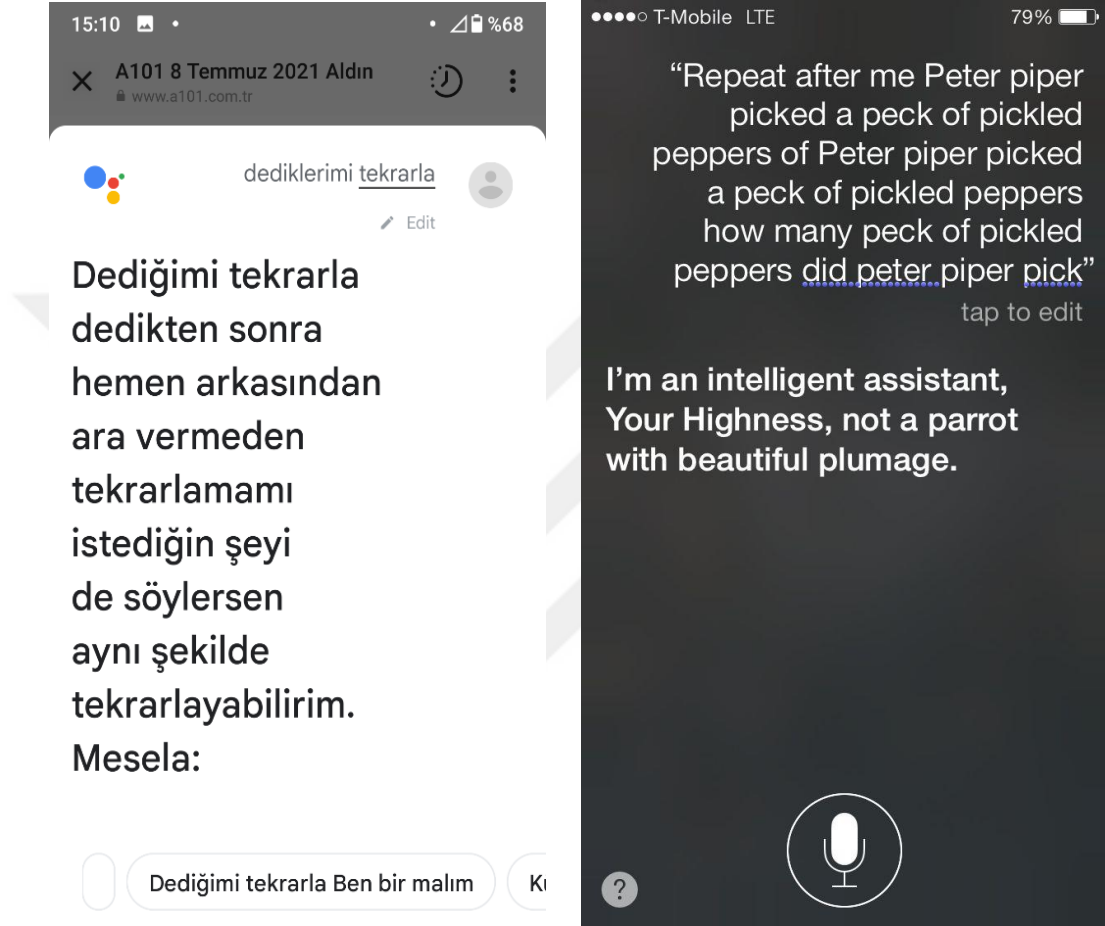


Figure 17. Siri has relatively superiority in terms of humour. However, Google Assistant has a limited and different approach to humour

Apart from these, user satisfaction is reduced and sometimes frustrating experiences are unveiled due to Siri's human-like character. Because of Siri's this feature, the mental model in the user mind and Siri's character are not matched. When a human-like approach from SMA is encountered by the user, SMA tends to be communicated just like a real person. However, Google Assistant is evaluated as less human-like. Therefore, more task-oriented behaviour, using simple commands, and simple words are more common among Google Assistant users. As a result, Google Assistant is much closer to the mental model on the user's mind. Consequently, the Google

Assistant experience resulted mostly positively than the Siri experience. Similarly, SMAs are mostly forgiven because of faulty results and expected modest performance by people who have more knowledge about technology than people do not have:

"I think I'm probably quite forgiving because I know how hard voice recognition is as a problem, and especially in the time frame that it's attempting to work on, just a few seconds, I know how difficult a problem that is. So it doesn't really annoy me if it doesn't work it out because I assume it's not going to because I know how difficult it is."

Rob (Luger, and Sellen, 2016)

In SMA interaction there are some resemblances. Recent studies showed that positive user experience was cultivated by not only correctness but also neutrality of responses. While the answer of Google Assistant 92% "good" and 97% correct, was found by participants. The response of Siri was much worse with 16% good and 14% correct. Besides female voices of Google Assistant were perceived as more natural and emotional than other assistants such as Siri, Alexa, and Microsoft Cortana (Berdasco , 2019). Nevertheless, Siri's voice was evaluated as natural by all of Siri users in this study with the latest updates. Even though there are significant differences between male-gendered voices and female-gendered voices, low-status language, and female-gendered voices have been an inclination to be preferred by users and developers (Habler, Schwind, and Henze, 2019). Preferring female voices was not complained about by both participants who have joined interviews and participants' opinions from literature reviews. It seems every participant is content with this decision. Clifford Nass from Stanford University said:

"It's much easier to find a female voice that everyone likes than a male voice that everyone likes. ... "It's a well-established phenomenon that the human brain is developed to like female voices." (Griggs, 2011)

It is related to combining a woman's voice with motherhood. Therefore, it felt more pleasing. Similarly, on the Moscow metro system, both male and female voices are used for the announcements. On the train, which goes to the city centre male-gendered voice is used (your boss calls you to work). However, in the train that crosses the city

centre and goes to the suburbs of Moscow, a female-gendered voice is chosen for the announcement (your wife calls you home) (alTerrAlexey, 2018). Additionally, more human-like features of SMA are required for a positive experience to be demonstrated on user recommendations. Even though a few of them emit they would be uncomfortable if SMA was more human-like, from a customization perspective, speech customization is located in user recommendations. It is related to SMA personality:

“If it’s possible to change mood [...] if I am a sarcastic person, I want her answers in the same sarcastic way or if I am a serious person, I want her answers to be in a very academic and serious way.” (Lopatovska et al., 2019).

There is some positive user opinion about Siri. If an attitude of SMA which is changed depending on the user's mood were provided, the user would be highly satisfied. A Siri user mentioned:

“There was one time I was very [sarcastic] to it, I was like ‘oh thanks that’s really helpful’ and it just said, I swear, in an equally sarcastic tone ‘that’s fine it’s my pleasure’”

Sarah (Luger, and Sellen, 2016)

Additionally, the personal and emotional touch by SMA is highly appreciated by the user and these lead to more personification. A study was conducted with two chatbots, Sarah and Caroline. While Sarah was human-like, Caroline was more robotic. When a human-like approach was observed by the user. More positive user experience was reported. In the study, a participant stated that having a baby. Sarah congratulated her. It was appreciated by the participant:

“It felt nice of her to care about me “ (Luger, and Sellen, 2016).

	Preferred chatbot		
	Sarah-Caroline	Caroline-Sarah	Preferred
P1		x	Sarah
P2		x	Sarah
P3		x	Sarah
P4	x		Caroline
P5	x		Sarah
P6	x		Sarah
P7	x		Caroline
P8	x		Caroline

Figure 18. The table shows preferred chatbots. Sarah is mostly preferred one. (Source: Luger, and Sellen, 2016)

“take care of yourself”. Siri tries to call the taxicab even further. A study that evaluated SMA’s as a whole showed that the most expected capabilities from SMAs are recognizing interruptions and having human-like behaviour. But there was a difference

Current Capabilities	FU (N = 53)	IU (N = 65)
	Med/M/SD	Med/M/SD
I need to speak differently	3/2.91/1	3/2.91/1.1
VPA misunderstands me (NLU)	2/3/1.41	2/3.09/1.3
VPA struggles with accent (ASR)	2/3/1.41	4/3.45/1.21
Results are irrelevant	4/3.36/1.36	2/2.82/1.17
Expected Capabilities	FU (N = 53)	IU (N = 65)
	Med/M/SD	Med/M/SD
Should Recognise Interruptions	5/4.45/0.87	4/4.09/0.83
Should be Human-Like	5/4.3/1.2	4/3.36/1.36
Should Have Personality	4/4.09/1.3	4/3.09/1.3
Should Ask More Questions	4/4.09/0.83	4/3.09/1.3

Figure 19. Rated current and expected capabilities of SMA by the frequent and infrequent users. (Source: Luger, and Sellen, 2016)

Moreover, it was demonstrated by the General depiction of the user satisfaction about them that Sarah was more preferable than Caroline (Figure 18). Sarah was preferred by participants because of these reasons: easy to communicate, existing typing awareness indicator, and funny conversions with a personal touch. A similar approach is observable in Google Assistant. When the user says: “I’m drunk”. It warns you:

between the needs of frequent users (FU) and infrequent users (IU). While all four criteria were thought equally important by infrequent users, being human-like and recognizing interruptions were more expected by frequent users (figure 19).

The other aspect of a personification of SMA is its capacity for using humour and irony. Studies show that in the case of humour usage in the technology, the personification of SMA is unveiled.

Under these circumstances, this topic is taken seriously by developers. Content writers from Pixar, and The Onion were hired by Google LLC for creating jokes and a conversational tone for Google Assistant (Lopatovska, 2019). SMA’s are developed by many corporations such as Apple, Google, or Microsoft. From a brand personification perspective, SMAs which add value to existing ways of interacting with a brand have a higher adaptation ratio by the user. As a result, it tends to be embraced by customers.

Value is added to a brand by SMA in these three ways:

- By helping the user to get a task in less time
- Providing richer interaction in the same amount of time.
- Assist the user to accomplish the same task in less time

When one of these situations is raised, brand impression increases significantly. Similarly, a participant who has taken part in an interview about Bank America's chatbot said:

“Bank of America’s Erica definitely improved my perception of the Bank of America brand. It is great to see my bank staying ahead of the curve.”
(AnswerLab, 2018).

Customization tends to be found distinctive by the user. Special assistant is desired in some situations. In this case, predicting frequently used locations without predefined inputs are desired from SMA by the user. Asking the user location of the job, or showing the irrelevant result is a frustrating experience. In this circumstance. User data is collected from Google databases by Google Assistant. User information is collected by Google not only from mobile devices but also from the personal computer. As a result, a more customized and user-oriented experience is generated by Google Assistant for the user.

There is much evidence that human-like SMA has a positive effect on the user. On the contrary, there are some opposing opinions about making Human-like SMA:

“Maybe we should stop trying to make Siri like a person because a person can only do so much. Make it better than a person, more helpful. Because if Siri's supposed to be a personal assistant, I would rather a personal assistant that can do way more than a human.” (Cowan et al., 2017).

4.3.2. Determining Differences of SMA with Considering Personification Side of Smart Mobile Assistants

Inanimate objects are sometimes personified by people. However, it is variable from user to user or objects to objects. For example, a car is loved or named by its users quite frequently. Nevertheless, some objects such as iron is rarely personified. In this chapter, the differences of SMA in terms of the personification side are evaluated. Are Google Assistants and Apple Siri personified by the user? If they are personified, which features are the causes of this SMA personification? Besides, the connection of user psychology is evaluated in this chapter.

SMA is mostly close to human assistant in the user mental model. That's why jokes, and a human-like approach of the user are monitored in human-to-SMA interaction. Human assistant is hired to assist one person. In these circumstances, it is personal. Therefore, it is customizable. During his/her service, lots of things and information are learned by the assistant. Then the service starts to be customized depending on the master's needs and desires. Consequently, customizability is demanded by many users in order to create a positive user experience. The feeling of users is positively affected by customizability. In addition, while interacting with SMA, predicting user needs, estimating the next event, and suggesting without asking are required features from human-like personal assistants. In Lopatovska's study, these abilities for developing more human-like SMA were recommended by many participants. The user considers SMA has personality, and feasible options for real-life assistance were indicated. Nonetheless, there is a dilemma in this case. Users are generally reluctant to share personal data and information. Besides, privacy is a common concern among many participants. However, information about its user is needed by SMA in order to identify its owner and to be personal. Moreover, 75% of consumers were ready to share their data with a company that they trust (Brill, Munoz, and Miller, 2019). Even all SMAs are developed and in service by trusted companies such as Apple, Google, Microsoft, and Amazon. The most important factor for data sharing by the user is data safety. However, Apple has a bad reputation because of hacked celebrity accounts in iCloud (Gorman, 2016).

Additionally, in the study, these assistants are described as polite for Siri, and not having the personality of Google Assistant by the user. Besides, as indicated in the study, personality is required (Lopatovska et al., 2019).

Personification is conducted by various strategies. One of them is brand personification is a character who personifies the brand (Cohen, 2013). This strategy is conveyed by brand messaging and other signifiers such as colour, dress and behaviour, etc... in these circumstances, google corporate colour scheme is used by Google Assistant in their symbol and icon. Similarly, the waving bar -which appears at the bottom of the screen while it is searching or processing- is designed based on Google corporate colour scheme. Additionally, its name “Google Assistant” is linked directly to its developer. Its activation command is “Hey Google”, which fosters brand identity in the user's mind. Similarly, personification of Google Assistant with Google LLC. is observed among participants in this study. Google is mostly known as a search engine among users. That’s why Google Assistant is perceived as a character who knows everything and is respectful. A participant stated that:

“When we do not know something or are not sure, we generally say let’s ask Hz. Google.” (Hz is the abbreviation of hazretleri which means his/her holiness)

(P7 A Siri user)

Nevertheless, brand personification is not demonstrated by Siri as much as Google Assistant. The stem of its name refers to its history that comes from Stanford Research Institute. Its name is irrelevant to other Apple products, such as iPhone, iMac, iPod, iWork, and iMovie. Moreover, feelings of being an ambassador of the brand about Siri are reported by none of the participants who have been interviewed.

Moreover, anthropomorphizing a brand has the potential for the user to identify connections to the brand. Anthropomorphism is where a brand or the product is called a name and dedicated characteristics of human beings such as Uncle Ben, mother Russia, or Mrs. Goodwrench (Brown, 2011). This generates the effect that “this brand is like me” or “I am like this brand” (Cohen, 2013). Additionally, brands as a partner and establishing a relationship dyad tends to be accepted by the user (Fleck, Géraldine,

and Valérie, 2014). Therefore, the connection between customer and brand is increased. However, it strictly depends on customer satisfaction and experience. Today, limited abilities, humour, and setting up unrealistic expectations, and lack of customization are common on both Siri and Google Assistant sides. Nevertheless, positive customer satisfaction causes personification. A participant that joined this study interview stated that:

“Sometimes she frustrates me, but I have a good relationship with her. I like her.”

P3 (A Google Assistant user)

Even though a helpful and friendly approach is reported for both Google Assistant and Apple Siri, there are subtle differences among them. While Siri is located on the impudent side, Google assistant is evaluated as having non-personality characteristics. It resembles an ice-cold librarian (AnswerLab, 2018). The human-like nature of Siri is appreciated by many participants. A participant said:

“She's always really sassy, like when she turns on my iPad, and she's like "Oh, what can I help you with?" [FG2, P2] “I think having a personality helps a lot in making it less than just like you're barking commands into your phone”. (Cowan et al., 2017)

Using Siri on multiple devices is another positive factor in user experience. While Google Assistant is available on just mobile devices and smart home speakers, Microsoft Cortana is usable just on PC. Siri can be interacted with on various devices such as iPhone, iPad, Mac, and smart home speaker. However, there are still a few users who disregard the personality of SMA and focus on tasks that SMA can accomplish:

“The tasks that I used it for, I didn't see a personality really in it.” (Cowan et al., 2017).

In contrast, some contradictory behaviour is observed by participants who focus on the task-oriented side of SMA and ignore its personality. This is supported by these words:

“I don’t care its personality. I focus on what task it can do. (.....) She makes me very angry due to foolish mistakes and misunderstandings. Sometimes, I insult her. She apologies in a polite way. It makes me upset.”

P3 (A Google Assistant User)



CHAPTER 5: DISCUSSION

Even though, a vast improvement in SMA technology, it is evaluated as primitive and unmatured technology these days. Until the technology is matured, it is preferred by early adopters who are mostly technology addicts, function-oriented. Besides, cost and maintenance are less considered by early adopters. Because of the high battery consumption, Apple Siri is switched off by some users. Moreover, functionality, hands-free interaction, and misunderstood commands are mostly discussed issues among participants.

According to user interviews in this study and literature reviews, the human tendency to personify inanimate objects is discussed in this study. Even though immature technology of SMA, personification is fostered by humour and emoji usage. Whereas emojis usage leads to personifications of Google Assistant, Apple Siri is personified by the users due to its sarcastic and humorous approach. However, the absence of humour can not be mentioned for Google Assistant. Even though it is limited according to Apple Siri, the existence of humour can be mentioned. When a sarcastic approach by SMA occurs, the reaction of the user is considerable.

According to user interviews, adoration to technology is reported only by Siri users, whereas sadness and happiness are reported by Google Assistant users. Besides, Google Assistant is judged for exploiting user emotion by its users. On the contrary, happiness and feeling special are reported by many Siri participants, according to the literature review. Nevertheless, in case considering the development of technology, the differences and similarities might be changed. According to the literature reviews, Google Assistant is criticized by some of the participants because of a lack of humour ability. However, this situation is tried to be changed by Google LLC. The content and joke writers are hired for writing jokes and content for Google Assistant. Similarly, according to interviewed users, the voice of siri was evaluated as more machine-like prior to the latest updates.

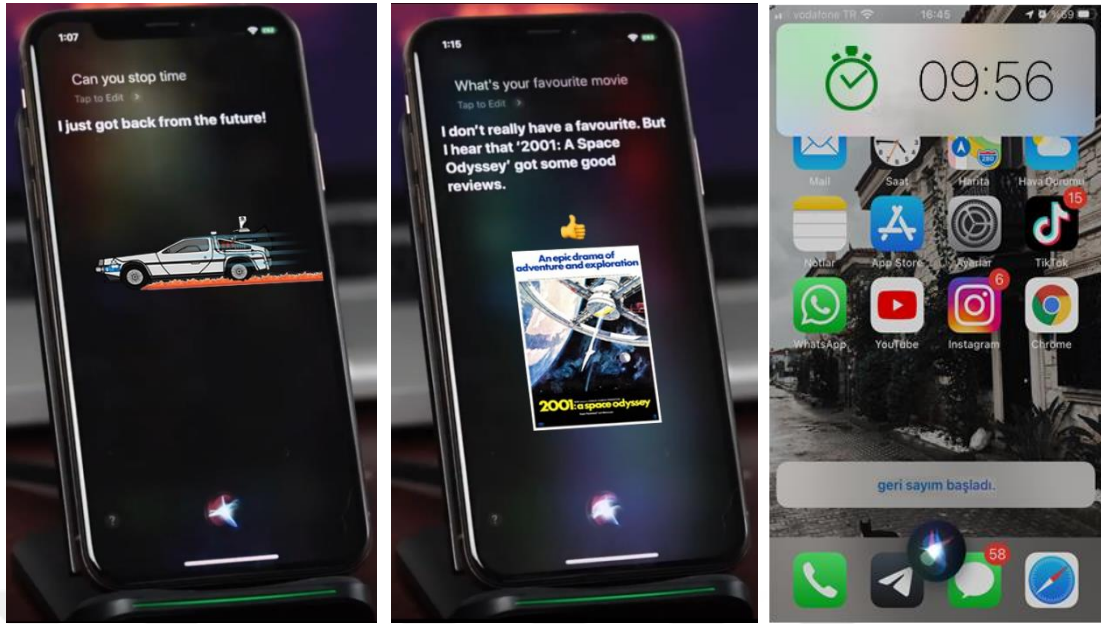


Figure 20. More vivid interaction is offered by Google Assistant. These figures were prepared for demonstrating richer multimedia interaction opportunities for Siri.

During interaction with the user, various multimedia notifications are used by SMA such as visual, auditory, and rarely vibration. Visual notification is not considered essential by some users, especially those who prefer to use SMA in driving. However, it is found essential by some participants because of fostering meaning and providing confirmation chances of SMA progression. SMA is still developing. Therefore, confirmation of words that are generated by the user is vital due to the detection of misunderstood commands. Additionally, while searching for something providing information both visually and auditory is appreciated by some participants. The most common interaction that is observed in human-to-SMA interaction is asking for the weather forecast. It is indicated in literature reviews that Siri used to launch a weather forecast application to show the weather forecast by saying "*Here is the weather*" when the weather forecast was demanded by the user. However, this is complained about by some users. Because the device may not be reached by the user. On the contrary, this issue was not reported by interviewed participants in this study, the weather forecast is declared by both Apple Siri and Google Assistant as both visual and auditory notifications. Beyond this, many more complaints about hands-free interaction while using SMAs are reported by interviewed participants.

5.1. Limitations

This study was conducted based on a qualitative research method. The feelings and experiences that are encountered by participants are collected and evaluated. However, every user is unique. Deeper knowledge and some different user opinions and user experiences can be diversified by expanding the participant pool. All interviewed participants are mentally and physically healthy. For people with disabilities, different experiences can be unveiled. Similarly, desiring SMA by a person with visually disabilities is mentioned in the interview in this study. SMA for people with disabilities can be discussed by further studies. Additionally, giving his ex-girlfriend a name is tried by one depressive Siri participant. The mental health of the participant was completely normal. However, SMA usage in different psychology situations can result in a different experience. SMA usage by mentally unhealthy people appeared to cause unpredictable results and some social implications. SMA usage by people who have various psychological and mental conditions can be discussed in further studies.

In addition, Amazon Alexa is another powerful counterpart. However, The Turkish language is not supported. Besides, Amazon Alexa is widely preferred as a smart home speaker by customers. Moreover, the smart home speakers market is a different topic of study. The study has an opportunity to expand to user experience on home smart speakers. Additionally, In case Alexa enters the smart mobile assistant market, this study might be re-conducted by evaluating Alexa as a new variable.

CHAPTER 6: FUTURE DESIGN OPPORTUNITIES

SMA are still developing. A great portion of negative experiences about human to SMA interactions that exist today will not be an issue in the near future. In interhuman interaction, facial expression and voice tone are a vital part of the act for people. Evaluating each other's psychology and feelings in interhuman communication are assisted by voice tones and facial expressions. However, the same feat cannot be achieved by today's SMAs. Human-like interactions would be developed by SMAs with gaining voice tone recognition and facial expression reading (AnswerLab, 2018). A participant stated:

“At a very high level, ability to identify meaningful data patterns that can surface preferences or problems and perhaps also predict future performance or outcomes.” (Dove et al., 2017).

Additionally, with the development of cybernetic technology, SMA can be interacted by humans via a chipset, which will be able to be implanted into the user body. In addition, calling on a smart assistant with a different name is discussed in John Scalzi's science fiction book "Old man's war." The relationship between John who is the main character of the book and the smart assistant that is implanted in his brain is mentioned. Similar to SMA this smart assistant is designed to accomplish some automated jobs such as, sending messages to recruits, downloading notifications, playing music or video, opening a document that is stored in the system. A different name from its default name is given to SMA by John. John chooses "asshole" to call on his smart assistant. Similarly, some names such as idiot, moron and asshole tried to be given to Google Assistant by one of the interviewed participants in this study. Additionally, calling friends and family members with a nickname is a common activity for this participant. Consequently, giving a nickname to SMA is impossible today. However, when it is possible, the personification of SMA will be fostered (Scalzi, 2005).

Additionally, SMAs suffer from a lack of intention recognition. Positive user experiences are led by estimating and understanding user needs, beliefs and psychology, therefore more psychological modes will be able to be integrated to SMA.

Different voice tones, vocabulary and more importantly multiple moods are demanded by some users today. A participant stated:

“If it’s possible to change mood [...] if I am a sarcastic person, I want her answers in the same sarcastic way or if I am a serious person, I want her answers to be in a very academic and serious way.” (Lopatovska, 2019).

Consequently, Apple Siri is evaluated as funny and sarcastic, more human-like by many users. That’s why, as juxtaposed to human assistants, it is closer to the user mental model. However, it is still just a software and because of its nature, some limitations exist. How to work voice recognition is not comprehended enough by many users and SMA tends to be interacted with like a real personal assistant. In these circumstances, limitations and weaknesses of SMAs are unveiled. Consequently, the negative experiences would be raised. Nevertheless, the abilities of SMA may be signaled to the user by using robot-like voice characteristics in order to manage the expectations of users. So, the tendency of using basic commands by the user would be prioritised. As a result, improving user experience is possible. Even though Siri has been found funnier and more human-like than Google Assistant, when the positivity and correcting of answers are evaluated, the performance of Google Assistant is reported higher by participants who use Google Assistant and graded higher by the users because of it.

The usage of SMAs is the main focus of this study. However, the interviews demonstrated that using SMA in a car is a very common activity. Hence, Smart Assistant in a car might be a logical design field. Besides, some car manufacturers are aware of this opportunity, and some primitive technologies are starting to be developed. Today, some technologies such as navigation or calling someone are preferred by the driver. Even though, sometimes it resulted in unpleasant and frustrating experiences (denge35, 2017). Similarly, a different driver experience based on touch screen and voice control for uncritical car instruments such as climate control, entertainment, mobile phone is provided by Ford U concept. Even though, instead of using just a voice control interface in their cars, the prototype is an implementation of a voice control interface that is combined with a graphical user interface (Kėpuska, and Bohouta, 2018).

Additionally, a cloud-based platform that allows humans to control a robot with gestures, natural speech-language, and body language is researched by the University of Southern California. However, this is a developer's platform. Speech and gesture-based applications that are generated by robots can be developed with the help of this platform. At the same time, non-expert users can interact with these robots through natural communications (Kėpuska, and Bohouta, 2018). This technological development fosters positive human to SMA interaction.

Moreover, some proposals are presented as a system. In one proposal, speech, image, video, touch, manual gestures, gaze, and head and body movements that are generated by the user are accepted as combined inputs. The existing structure of smart assistants is proposed to modify according to these new requirements and input models to design a new smart assistant model. These are called ASR Model, Gesture Model, Graph Model, Interaction Model, User Model, Input Model, Output Model, Inference Engine, Cloud Servers, and Knowledge Base. Even though most of the models are related to technical parts of the SMA development, some of them are crucial for designing positive SMA interaction. These are knowledge base, graph model, gesture model, ASR model, and user model. A knowledge base is composed of the user's gestures, voice, language, body, and personal information. In the graph model, images and videos are analyzed in real-time. In the gesture model, body movement, gestures, and facial expression is read and sent to a specialized server. the server interprets these inputs and sends them to the device to perform the results. The utterance and voice are recorded via the microphone and converted to text by the server in the ASR model. The user's information and personal needs and desires are kept in the system in order to make a reliable decision as to the main purpose of the user model. Besides, this model was tested in Google cloud services and Amazon web services. According to test results, whole concepts in this system are the best choice for designing and developing next-generation virtual assistants following some hardware and software improvement. The result of the study shows that user to SMA interaction could be increased and affected positively by using these technologies, such as gesture recognition, image/video recognition, speech recognition, and the Knowledge Base (Kėpuska, and Bohouta, 2018).

Today, health data is collected by smart wearable devices. These are generally processed for sport activities. However, this data carries valuable information about user psychology. As it is discussed above, SMA suffers from a lack of user facial expression and gestures in order to estimate user psychology. The data such as blood pressure and heart rate can give clues about user psychology and a chance to estimate user needs and desires.

In Addition, gender differences among users are considerable for enhancing user experience. Interests and requirements of male and female users are different. Therefore, SMA can switch its mood and behaviours depending on users' genders.

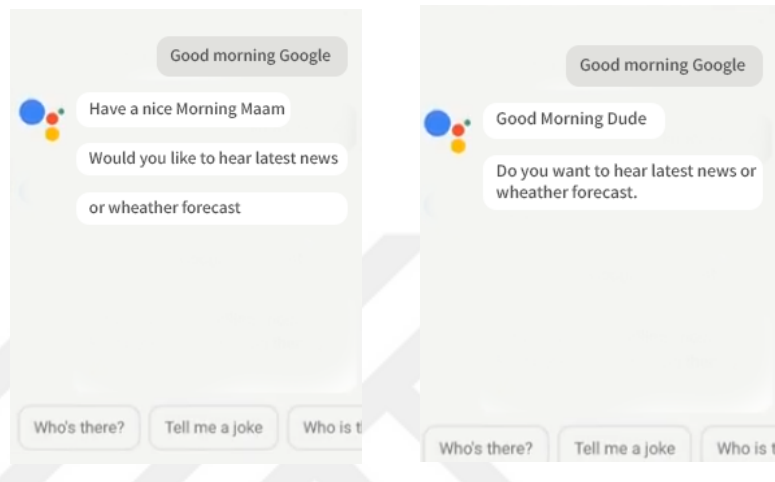


Figure 21. These figures were prepared for demonstrating responses of SMA that estimates different types of users' requirements and needs.

SMA usage in driving is reported by most of the interviewed participants. Hands-free interaction opportunity is the main motivation factor among participants. For enhancing user experience, avoiding interruption of hands-free interaction generates positive results. Additionally, driving sometimes can be a stressful task. The needs and psychology of drivers differ from traffic jams to traveling to the seaside. If the psychology of the driver can be monitored by smart wearable devices, the mood of SMA can be changed. In addition, the location of the driver and information about the traffic that is collected by navigation software can give a clue about driver psychology and mood. According to these data, the recommendations and behaviour of SMA can be determined. Moreover, even though SMA-to-car connection is offered by a few car manufacturers, SMA can interact with the driver visually via a Lcd screen on the dashboard or head-up display without distracting the driver. SMA is preferred by the users for navigation purposes. However, data from sensors of the car can be monitored and interpreted by the SMA. Therefore useful recommendations about the car can be

offered by SMA such as clues about reducing fuel consumption and safe driving. Besides, some of the non-critical equipment can be controlled by SMA such as climate control and entertainment.



CHAPTER 7: CONCLUSION

With the launch of Siri by Apple in 2011, the human-to-computer interaction has gained a different dimension. Even though talking and personality-owned computers are depicted in science fiction from Enterprise computer in Star Trek, to the KITT in the Knight Rider tv series. It has been out of reach of consumers until today. Basically, SMAs are software. However, most of the users of SMAs are not software engineers. Therefore, the represented model has gained importance in order to be useful and desirable and to generate a positive user experience.

After one year of launching Apple Siri, a rival which is named Google Now was released by Google LLC. Voice search was a fundamental feature of Google Assistant. It was built for more functional purposes rather than Apple Siri. Then it has been evolved to Google Assistant. Task-oriented features carry on.

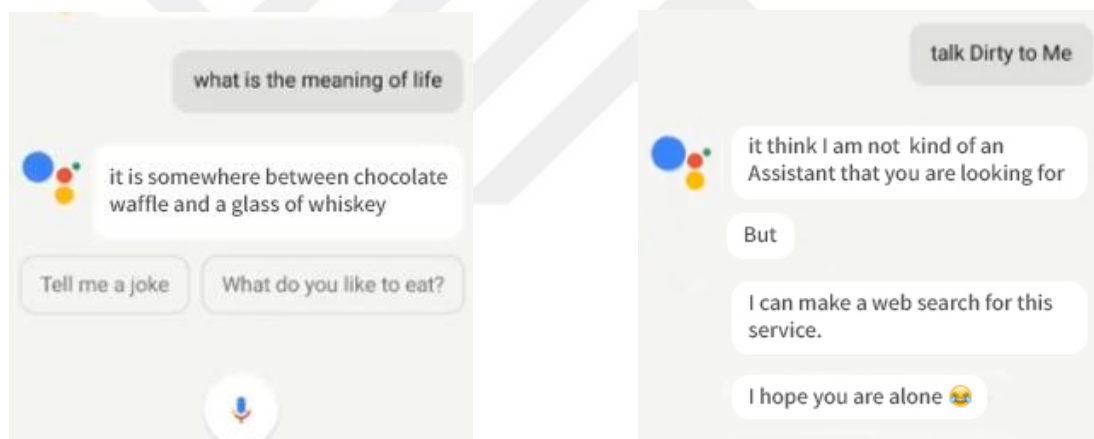


Figure 22. Google Assistant is evaluated as ice-cold-librarian. these figures were prepared based on the user expectations

According to literature reviews, Siri is evaluated as more sarcastic and funny approach, whereas Google Assistant is described as an ice-cold-librarian.

However, some emotional approach was reported by one interviewed Google Assistant participant. Even though Google Assistant is mostly described as having no personality, this trait can be considered as a personality in itself. As a multimedia notification, the voice of SMA was mostly liked by the users personally. Nevertheless, the voice of Siri is evaluated as more human-like than Google Assistant. In addition,

the Robot-like voice of Google Assistant was appreciated by one of the participants due to indicating the development of technology. Similarly, instead of developing human-like featured SMA, exploiting the SMAs in their own opportunities would provide a positive experience for some users. The satisfaction rate among the interviewed participants is above average. Additionally, considering the average experience of users on SMA interaction, the basic and needed or desired actions that are requested by the users are performed by SMA adequately. The SMAs are used in various places for various purposes. The primary motivation beyond the SMA interaction is time-saving. Adding appointments using SMA is by far quicker and more simple than doing it by hand. SMAs are preferred by most of the participants while driving. Driving is a job that has to be done carefully. That's why SMA is described as a lifesaver by participants who prefer SMA in driving. However, because of the loss of visual contact, visual notifications of SMA are easy to ignore. However, auditory confirmation is more detectable. Therefore, in terms of confirmation, auditory confirmation is the most effective one. Similarly, just auditory and text-based notifications were enough for interaction by seven of ten interviewed participants. visual and auditory notifications are not used by home smart assistants even further.

As an auditory notification, the voices of all SMA that are used by interviewed participants are female. Female voices are stated as more natural and pleasing by participants both at literature reviews and conducted interviews.

The commands that are given by the user and the responses that are provided by SMA are confirmed both visually as text and auditory as voice. However, Google Assistant and Apple Siri differ in that point. The background colour and text colour are completely different from each other. The most readable preferences are chosen by Google Assistant as dark text colour on white background. Additionally, Google Assistant is differentiated by using emojis and stickers. This design choice is indicated by many participants as chatting like a real person. Besides, Emoji usage is appreciated by participants. Moreover, more emoji usage is requested by one participant. The Emoji usage of Google Assistant is based on mostly expressing its own emotional situation and sometimes fostering existing situations. Google Assistant tends to be personified by the user due to this ability.

While listening to the user, progress is displayed on the screen by both Google Assistant and Apple Siri. While the weaving bar is displayed on the screen by Apple Siri, a colourful stick is demonstrated by Google Assistant. On the contrary, the colourful stick that is displayed by Google Assistant is designed based on Google corporate identity colour scheme. Besides, those colours are used for designing the symbol of Google Assistant.

In order to wake up Google Assistant, the "Hey Google " order is used by the user. As a result, Google Assistant is more brand personification oriented than Apple Siri, whereas Siri is evaluated as having its own personality regardless of its developer. However, due to its sarcastic and funny approach, Apple Siri's personality is more remarkable than Apple's own brand identity. Forcing the user to use the developers' application by SMA is reported by participants that are indicated in literature reviews. However, it is not reported by any of the interviewed participants in this study. Besides, any negative experiences about it were not reported.

SMA usage is described by some technology experts as using a computer in its dark ages in the 1970s. At that age, the commands have to be memorized by the user. Similarly, there are some similarities in today's SMA interaction. Human to SMA interaction is evaluated negatively in case forming sentences in a different way. To combat this issue, the hints are displayed by Google Assistant on its interface. Nevertheless, the same aid is not offered by Apple Siri. This is offered as a suggestion by an interviewed Apple Siri user in this study even further.

Google Assistant and Apple Siri seem to be competitors of each other in the market. However, each SMA comes forward with some features. Even though various multimedia notifications are used for both Google Assistant and Siri, Google Assistant is evaluated as offering more vivid interaction because of emoji and sticker usage, and clarity on some visual notifications. As auditory notification is mostly based on voice, Apple Siri is found more vivid and human-like due to its natural voice. However, Apple Siri is only equipped with this voice characteristic with the latest updates that are indicated by interviewed participants. Characteristics of both Google Assistant and Apple Siri are different. While Apple Siri is evaluated as more sarcastic and funny compared to Google Assistant, Google Assistant comes forward with task-oriented

features and having the ice-cold characteristic. Moreover, although it is limited, humour ability exists in Google Assistant interaction. Inanimate objects tend to be personified by a human. This is a legacy that is left by the ancestors of homo sapiens. Consequently, SMA tends to be personified by the user. However, the personification of SMA is unveiled differently. When Siri is personified due to its sarcastic and funny characteristics, Google Assistant is personified because of texting like a real person, emoji usage and bonding with its developer's company. SMA technology has great potential. As it is discussed in the "Future Opportunities" section, human-to-SMA interaction might gain sophisticated experience for the users when SMAs are equipped with implant technology and some technologies, such as more human-like voice tone, gesture and facial expression recognition.



REFERENCES

- [alTerrAlexey]. (2018). Информатор: Сокольническая линия (старое). [Online] Available at:
https://www.youtube.com/watch?v=obk3x2UpXM8&list=PLQ3ZNQGki2tnXreQ_5Uv7v5ligmrk5BYP (Accessed 11 May 2021).
- AnswerLab. (2018). *Elements of a Successful Digital Assistant and How to Apply Them* [Online]. Available at: <https://cdn2.hubspot.net/hubfs/637744/AnswerLab%20%20Successful%20Digital%20Assistants.pdf> (Accessed 10 April 2020).
- Berdasco, A., Gustavo, L., Diaz, I., Quesada, L., Guerrero. L. A., *User Experience Comparison of Intelligent Personal Assistants: Alexa, Google Assistant, Siri and Cortana*. University of Costa Rica. San José. 2-5 December 2019.
- Bernard, M. L., Chaparro, B. S., Russell, M. (2001) *Examining Automatic Text Presentation for Small Screens: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Vol.45(6)*, pp. 637–639.
- Brill, T., Munoz, L. and Miller, R. J. (2019) *Siri, Alexa, and Other Digital Assistants: A Study of Customer Satisfaction with Artificial Intelligence Applications*, *Journal of Marketing Management*, Vol. 35(15-16), pp. 1401-1436.
- Brown, S. (2011) *It is Alive Inside! A Note on the Prevalence of Notification*, *Irish Marketing Review*, Vol. 21, pp. 4-5.
- Budiu, R. and Laubheimer, P. (2018) *Intelligent Assistants Have Poor Usability: A User Study of Alexa, Google Assistant, and Siri*, *Nielsen Norman Group*, 22 July [Online]. Available at: <https://www.nngroup.com/articles/intelligent-assistant-usability/> (Accessed 29 march 2021).
- Cohen, R. J. (2013) *Brand Personification: Introduction: Introduction and Overview*, *Psychology & Marketing*, Vol. 31(1), pp. 1-30.
- Cooper, A., Reimann, R. and Cronin D. (2007) *Implementation Models and Mental Models*. In: *About Face 3 The Essentials of Interaction Design*. Indianapolis: Wiley Publishing, Inc., pp. 27-37.

Cowan, B. R., Pantidi, N., Coyle, D., Morrissey, K., Clarke, P., Al-Shehri, S., Earley, D., Bandeira, N. “*What Can I Help You With?* ”: *Infrequent Users*’. *MobileHCI '17: Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*. Vienna. September 2017.

[denge35]. (2017). Türk'ün Navigasyonla İmtihanı. [Online].

Available at: <https://www.youtube.com/watch?v=QFcBHvU2Ekk> (Accessed 16 July 2021).

Dillow, C. (2011). Apple's New iPhone Assistant is a Voice-Activated DARPA Spinoff, *Popular Science*, 6 October [Online]. Available at: <https://www.popsci.com/technology/article/2011-10/apples-new-iphone-assistant-voice-activated-darpa-spin-out/> (Accessed 18 November 2020).

Dove, G., Halskov, K., Forlizzi, J., Zimmerman, J. *UX Design Innovation: Challenges for Working with Machine Learning as a Design Material: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, Denver. 2017.

Fleck, N., Géraldine, M. and Valérie, Z. (2014) *Brand Personification through the Use of Spokespeople: An Exploratory Study of Ordinary Employees, CEOs, and Celebrities Featured in Advertising*, *Journal of Psychology & Marketing*, Vol. 31(1), pp. 84-92.

Gorman, S. (2016). The man accused of hacking into celebrity iCloud accounts and stealing naked photos has agreed to plead guilty, *Insider* 2 July [Online]. Available at: <https://www.businessinsider.com/r-illinois-man-agrees-to-plead-guilty-in-celebrity-hacking-case-2016-7> (Accessed 25 June 2021).

Griggs, B. (2011). Why computer voices are mostly female, *CNN Business*, 21 October [Online]. Available at: <https://edition.cnn.com/2011/10/21/tech/innovation/female-computer-voices/index.html> (Accessed 9 June 2021).

Guerrero, L. K., and Floyd, K. (2008) *I. Introduction*. In: *Nonverbal Communication in Close Relationships*. New Jersey: Lawrence Erlbaum Associates, Inc., pp. 1-10.

Habler, F., Schwind, V., Henze, N. *Effects of Smart Virtual Assistants' Gender and Language: Proceedings of Mensch Und Computer 2019*. Association for Computing Machinery. 08 September 2019.

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

Këpuska, V., and Bohouta, G. *Next-Generation of Virtual Personal Assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home): Proceeding of IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*. Las Vegas. 2018.

Klein, A. M., Hinderks A., Schrepp M., Thomaschewski J. *Measuring User Experience Quality of Voice Assistants. Proceedings of 15th Iberian Conference on Information Systems and Technologies (CISTI)*. Seville, Spain. 24-27 June 2020.

Kolko, J. (2010) In: j. Kolko, ed. *Thoughts on Interactions Design*. Burlington: Elsevier, pp. 11-42.

Krazit, T. (2010). Google finding its voice, *Cnet*, 31 August [Online]. Available at: <https://www.cnet.com/news/google-finding-its-voice/> (Accessed 16 November 2020).

Langer, M. (2012) *Dragon Dictate 2.5: Visual QuickStart Guide*. 1st ed. California: Peachpit Press.

'Language Support in Voice Assistants Compared'. (2021). *Summalinguae*, 21 April [Online]. Available at: <https://www.globalme.net/blog/language-support-voice-assistants-compared/> (Accessed 19 November 2020).

Lopatovska, I. (2019) *Classification of humorous interactions with intelligent personal assistants*, *Journal of Librarianship and Information Science*, Vol. 52(9), p. 2-4.

Lopatovska, I., Griffin, A. L., Gallagher, K., Ballingall. C., Rock. C., and Velazquez. M. (2019) *User Recommendations for Intelligent Personal Assistant*, *Journal of Librarianship and Information Science*, Vol. 52(2), pp. 577-591.

Lowerre, B. T. (1976) *The HARPY Speech Recognition System*. Pittsburg: Department of Computer Science Carnegie-Mellon University.

Luger, E., Sellen A. *"Like Having a Really Bad PA": The Gulf between User Expectation and Experience of Conversational Agents: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York. May 2016.

Mikkelsen, D. (2010). General Motors Replies to Bill Gates. Snopes [Online]. Available at: <https://www.snopes.com/fact-check/car-balk/> (Accessed 4 July 2021).

Nielsen, J. (1994). 10 Usability Heuristics for User Interface Design, *Nielsen Norman Group*, 24 April [Online]. Available at: <https://www.nngroup.com/articles/ten-usability-heuristics/> (Accessed 10 June 2020).

Norman, D. A. (1999) *Growing up: Moving from Technology-Centered to Human-Centered Products*. In: *The Invisible Computer*. London: MIT Press, pp. 24-49.

Öquist, G., Lundin, K. *Eye Movement Study of Reading Text on a Mobile Phone Using Paging, Scrolling, Leading, and RSVP: Proceedings of the 6th International Conference on Mobile and Ubiquitous Multimedia*. Association for Computing Machinery, Oulu. 12 December 2007.

Pieraccini, R. (2012) *The Voice in the Machine: Building Computers That Understand Speech*. 1st ed. London: The MIT Press.

Sauro, J., and Lewis, J. (2012) *Chapter 2 - Quantifying User Research*,. In: M. Kaufmann, ed. *Quantifying the User Experience*. 2nd ed. Cambridge: Morgan Kaufmann, pp. 9-18.

Scalzi, J. (2005) *Old Men's War*. Tor Books, pp. 97-101.

Schumacher, R. M. (2009) *Chapter 1 - Foundations and definition*. In: *Handbook of Global User Research*. 1st ed. Burlington: Morgan Kaufmann, pp. 1-20.

Science: 'Shoebbox Is Listening' (1961). *The Time*, 24 November [Online]. Available at: <http://content.time.com/time/subscriber/article/0,33009,828878,00.html> (Accessed 26 September 2021).

- Sharmin, S., Špakov, O. and Rähä, K.-J. (2012) *The Effect of Different Text Presentation Formats on Eye Movement Metrics in Reading*, Journal of Eye Movement Research, Vol. 5(3), pp. 1-9.
- Tang, Y. and Hew, K. F. (2019) *Emoticon, Emoji, and Sticker Use in Computer-Mediated Communication: A Review of Theories and Research Findings*, International Journal of Communication, Vol. 13, pp. 2457-2483.
- Thomson, D., and Filik, R. (2016) *Sarcasm in Written Communication: Emoticons are Efficient Markers of Intention*, Journal of Computer-Mediated Communication, Vol. 21(2), pp. 105-120.
- Tilman, M. (2021). What is Apple Face ID and how does it work?, *Pocket-lint*, 3 September [Online]. Available at: <https://www.pocket-lint.com/phones/news/apple/142207-what-is-apple-face-id-and-how-does-it-work> (Accessed 09 may 2021).
- Tulshan, A. S., Dhage, S. N. *Survey on Virtual Assistant: Google Assistant, Siri, Cortana, Alexa. 4th International Symposium SIRS 2018*, Sardar Patel Institute of Technology, Mumbai, 19–22 September 2018.
- Wang, A-H., Fang J-J., Chen C-H.(2003) *Effects of VDT Leading-Display Design on Visual Performance of Users in Handling Static and Dynamic Display Information Dual-Tasks*. International Journal of Industrial Ergonomics. Vol. 32(2), pp. 93-104.
- Wiggers, K. (2019). Which voice assistant speaks the most languages, and why ?, *Venturebeat*, 2 February [Online]. Available at: <https://venturebeat.com/2019/02/02/which-voice-assistant-speaks-the-most-languages-and-why/> (Accessed 11 04 2021).
- Zhou, R., Hentschel, J. and Kumar, N. *Goodbye Text, Hello Emoji: Mobile Communication on WeChat in China. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, Denver. 02 May 2017, pp. 748-759.