ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE ENGINEERING AND TECHNOLOGY

PERSONALITY IN CONVERSATIONAL USER INTERFACES: EXTROVERTED AND INTROVERTED CHATBOTS

M.Sc. THESIS

Gonca İrem ÇAKIROĞLU

Department of Industrial Product Design

Industrial Product Design Programme

JUNE 2019

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE ENGINEERING AND TECHNOLOGY

PERSONALITY IN CONVERSATIONAL USER INTERFACES: EXTROVERTED AND INTROVERTED CHATBOTS

M.Sc. THESIS

Gonca İrem ÇAKIROĞLU
(502151923)

Department of Industrial Product Design

Industrial Product Design Programme

Thesis Advisor: Assoc. Prof. Dr. Çiğdem KAYA PAZARBAŞI

JUNE 2019

ISTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

SOHBETE DAYALI KULLANICI ARA YÜZLERİNDE KİŞİLİK: DIŞA VE İÇE DÖNÜK SOHBET BOTLARI

YÜKSEK LİSANS TEZİ

Gonca İrem ÇAKIROĞLU
(502151923)

Endüstri Ürünleri Tasarımı Anabilim Dalı

Endüstri Ürünleri Tasarımı Programı

Tez Danışmanı: Doç. Dr. Çiğdem KAYA PAZARBAŞI

HAZİRAN 2019

Gonca İrem ÇAKIROĞLU, a M.Sc. student of İTU Graduate School of Science Engineering and Technology student ID 502151923, successfully defended the thesis/dissertation entitled “PERSONALITY IN CONVERSATIONAL USER INTERFACES: EXTROVERTED AND INTROVERTED CHATBOTS”, which she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor : **Assoc.Prof. Dr. Çiğdem KAYA PAZARBAŞI** ....
İstanbul Technical University

Jury Members : **Prof. Dr. Gülname TURAN** ....
Istanbul Technical University

Assoc. Prof. Dr. Ayşe Elif ÇOŞKUN ORLANDI ....
Kadir Has University

Date of Submission : 03 May 2019
Date of Defense : 13 June 2019

To my family and friends,

FOREWORD

I would like to thank my supervisor, Cigdem Kaya Pazarbasi, for her guidance. I thank Burcu Kartal for her excellent statistical analysis skills, and I am grateful to my sister for her support in the thesis. I would like to thank Garanti BBVA Experience Design and JamesBot Teams for their support. Also, I thank my family and friends for their patience. I thank Dilara Yesilova for sharing this adventure with me. Finally, I would like to thank Muhammet Ramoglu for his tremendous guidance in the thesis. He always believed in me, which helped me complete this journey.

June 2019 Irem Cakiroglu
(Designer)

TABLE OF CONTENTS

Page

ABBREVIATIONS

CUI	**:** Conversational User Interface
EC	**:** Extroverted Chatbot
EP	**:** Extroverted Participant
HCI	**:** Human-Computer Interaction
HRI	**:** Human-Robot Interaction
IC	**:** Introverted Chatbot
IP	**:** Introverted Participant
VPA	**:** Virtual Personal Assistant

LIST OF TABLES

Page

LIST OF FIGURES

Page

PERSONALITY IN CONVERSATIONAL USER INTERFACES: EXTROVERTED AND INTROVERTED CHATBOTS

SUMMARY

Conversational user interfaces have become a part of our lives today. Chatbot is one of the conversational user interfaces that many companies have been using them as their own virtual personal assistant. One of the reasons for this is to provide customers with 24/7 service, low cost compared to customer agents and easy access to customers's personal data history. Furthermore, chatbots have become an attractive tool for users because they provide support to users with quick access to the answers they are looking for, eliminating the need for additional applications, and allowing them to save their conversations. In other words, it is benefitical for both users and companies. Therefore, in the face of the rise of chatbots, designers' mission should be to represent users a better experience.

One of the critical points of establishing better experiences and relationships between chatbots and users is to design qualified interactions in chatbots. For this reason, the questions that designers and design researchers should ask themselves are as follows: What should be considered when creating a chatbot as a designer? How should a relationship be established with users? What is the importance of personality in these products that behave lifelike? This research aims to answer these questions and to investigate the personality creation in text-based user interfaces to improve user experience in products that behave like living things. For this purpose, an online study has been conducted on chatbots, which are also described as text-based personal assistants. The aim of the study is to investigate whether users will be able to recognize designed chatbot personalities, whether personalities of chatbots will affect perception of chatbots' capability, and whether personalities of the users will affect their perception.

For this reason, two chatbots with different personalities were designed. Two opposite personalites have been used in the study as extroverted and introverted. Seven important parameters were identified to design these personalities. These can be listed as the *tone of voice, avatar, expressions, use of emojis, use of buttons, the amount and frequency of the answers*, and *willing to have a conversation*. Based on these items, chatbots were designed and scenarios were created. The two chatbots performed the same tasks according to their personalities. Participants who saw conversation of chatbot in a video responded to a survey and perception of these chatbots were measured as whether they were perceived as extroverted or introverted as they were attributed by the users, how their capabilities were perceived, whether users' personalities affect their perception on *fun, competence, usefulness of interaction*, and *liking* dimensions. In summary, the aim of the thesis is to examine the importance of personality in text-based user interfaces and guide designers to create better chatbot experience for users.

SOHBETE DAYALI KULLANICI ARA YÜZLERİNDE KİŞİLİK: DIŞA VE İÇE DÖNÜK SOHBET BOTLARI

ÖZET

Sohbet tabanlı kullanıcı ara yüzleri bugün hayatımızın bir parçası haline geldi. Birçok şirket kendi kişisel asistanını yaratmaya başladı. Bunların başında da metin tabanlı kişisel asistan olarak da adlandırabileceğimiz sohbet botları gelmektedir. Bu alandaki yatırımların yükselişinin nedenlerinin başında müşterilere 7/24 hizmet sağlamak, canlı müşteri temsilcilerine kıyasla düşük maliyet elde etmek ve müşterilerinin kişisel verilerine kolay erişim gibi avantajlar gelmektedir. Dahası, sohbet botları kullanıcıların aradıkları yanıtlara hızlı erişim sağladıkları, ek uygulama ihtiyacını ortadan kaldırdıkları ve konuşmalarını saklama imkanı sundukları için kullanıcılar tarafından da çekici bir araç haline gelmişlerdir. Karşılıklı yarar sağlayan bu aracın günlük hayatımızda yer bulmaması elde değildir. Bu nedenle, sohbet botlarının kullanımının yükselişi karşısında tasarımcıların görevi, kullanıcılara daha iyi bir deneyim sunmak olmalıdır.

Kullanıcı ara yüzleri iki alt başlık altında incelenebilir. Bunlar, metin tabanlı ve ses tabanlı kullanıcı ara yüzleridir. Ses tabanlı kullanıcı ara yüzleri genellikle sanal kişisel asistan olarak da adlandırılıp sohbet botlarına göre daha akıllı bir bot oldukları algısı yaratılmaktadır. Ancak botların yetkinlikleri kullandıkları araçtan ayrı olarak kullandıkları alt yapıya göre ayrışmaktadır. Bun nedenle, bu tezde, sanal asistanlar hem metin hem de ses tabanlı ara yüzleri adlandırmak için kullanılmıştır. Sohbet botları ise metin tabanlı kişisel asistanlar anlamında kullanılmaktadır.

Metin tabanlı bu kişisel asistanlar ile kullanıcılar arasında daha iyi ilişkiler kurmanın kritik noktalarından biri ise nitelikli sohbet botları tasarlamak olarak düşünülebilir. Bu nedenle, tasarımcıların ve tasarım araştırmacılarının bu konuda kendilerine sorması gereken sorular şu şekilde sıralanabilir: Tasarımcı olarak sohbet oluştururken nelere dikkat edilmeli? Kullanıcılar ile nasıl bir ilişki kurulması gerekir?

Sohbet botlarında kişiliklerin tasarlanması potensiyeli olan bir alandır. Birçok kullanıcı günlük hayatında karşılaştığı bu sohbet botlarına istemsizce dahi olsa bir kişilik atfetmektedirler. Çeşitli araştırmalar, sohbete dayalı kullanıcı ara yüzlerinde kişiliğin kullanıcılarda güven duygusu oluşturmada kritik bir nokta olduğunu vurgulamaktadır. Sohbete dayalı kullanıcı ara yüzlerinde kişilik kullanımı ile güvene dayalı bir ilişki kurmanın mümkün olabileceği iddia edilmekte olup kullanıcı ile ürün arasındaki iletişimi güçlendirdiği savunulmaktadır. Ayrıca, oluşturulan bu güven sayesinde kullanıcı ve ürün arasındaki bağın kurulmasına katkıda bulunabilineceği de söylenebilir. Aynı zamanda bu bağ, ürün ve markanın kullanıcı ile olan duygusal etkileşimini de arttırabilir.

Bu tezde, sohbet botlarında kişiliğin araştırılmasının bir başka nedeni ise kullanıcı ile ürün arasındaki duygusal bağı güçlendirerek kullanıcı ile olan etkileşimi arttırmanın yollarını aramaktır. Günümüzde teknolojisi ile birlikte ürünler üzerindeki algı ve beklenti de değişmiştir. Bugün, ürünlerde sadece teknik açıdan fonsiyonellik değil

aynı zamanda kullanıcıların duygularına da hitap eden ürün tasarımları söz konusu olmuştur ve bunlar da artık ürünlerin bir fonksiyonu olarak adlandırılmaktadır. Sohbete dayalı kullanıcı ara yüzlerine kişiliğin kazandırılması, ürün ile kullanıcı arasındaki duygusal bağın yaratılmasında bir yöntem olarak kullanılabilir. Bu nedenle, bu konuyla ilgili araştırmaların sohbete dayalı kullanıcı ara yüzlerinin tasarımına önemli katkı sağlayacağı düşünülmektedir.

Bu tez araştırması, tasarımcılara yol göstermek amacıyla kullanıcı ile duygusal bir bağ kurmanın yollarını metin tabanlı kullanıcı ara yüzlerinde kişilik yaratımı ile araştırmayı hedeflemektedir. Bu amaçla, metin tabanlı kişisel asistanlar olarak da tarif edilen sohbet botları üzerinden bir çalışma yapılmıştır. Çalışmanın amacı, sohbet botlarına atfedilen kişiliklerin kullanıcılar tarafından nasıl anlaşıldığı, sohbet botlarının kapasitesinin algılanmasında kişiliklerinin etkili olup olmadığını ve kullanıcıların kişiliğinin bu algıları etkileyip etkilemeyeceğini araştırmaktır.

Bu amaç doğrultusunda oluşturulan hipotezler ve hipotezlere ait araştırma soruları şu şekilde sıralanabilir;

Hipotez 1: Katılımcılar sohbet botlarının kişiliklerini tanıyabilecektir.

Katılımcılar dışa dönük sohbet botunu dışa dönük, içe dönük chatbot'a içe dönük olarak mı düşünüyorlar?

Dışa dönük ve içe dönük sohbet botları için tasarım kriterleri nelerdir?

Hipotez 2: Sohbet botlarının kişilikleri, botların yetkinliğinin algılanmasında etkili olmayacaktır.

Sohbetler arasında kabiliyet ve etkileşimin yararı boyutlarında önemli bir fark var mı?

Hipotez 3: Katılımcıların kişilikleri, farklı kişilikleri olan sohbet botlarının nasıl algıladıklarını etkileyecektir.

Katılımcıların kişilikleri, chatbot kişiliğinin algılanmasını etkiler mi?

Katılımcıların kişilikleri; sohbet botlarının eğlence, kabiliyet, etkileşimin yararı ve sevme boyutlarını etkiliyor mu?

Araştırma kapsamında farklı kişilikleri olan iki sohbet botu tasarlandı. Bu kişilikler için biri dışa dönük, biri içe dönük olmak üzere zıt kutuptaki iki kişilik faktörleri tercih edildi. Bu kişilikleri tasarlamak için 7 önemli unsur belirlendi. Bunlar; sohbet botunun *konuşma şekli, avatarı, ifadeleri, emoji kullanımı, buton kullanımı, cevapların miktarı ile sıklığı* ve *konuşma isteği* olarak listelenebilir. Bu öğelere göre sohbet botları tasarlandı ve senaryolar oluşturuldu. İki sohbet botu da aynı görevleri kendilerine atfedilen kişilikleriyle gerçekleştirdiler. Böylece, aynı görevi kendi kişiliğine göre gerçekleştiren bu sohbet botlarının kullanıcılar tarafından da atfedildiği gibi dışa dönük veya içe dönük olarak algılanıp algılanılmadığı, kapasitesilerinin nasıl algılandığı ve kullanıcıların kişiliklerine göre sohbet botu hakkındaki algılarının değişip değişmediği *eğlence, kabiliyet, etkileşimin yararı* ve *sevme* boyutları ile ölçümlendi.

Araştırmanın sonuçları katılımcıların içe dönük sohbet botunu dışa dönük, dışa dönük sohbet botunu ise daha dışa dönük olarak algıladıklarını göstermektedir. Başka bir deyiş ile, içe dönük olarak tasarlanan sohbet botu dışa dönük olarak tasarlanana göre daha içe dönük algılanmıştır. Bu durumda 1. Hipotez doğrulanmıştır. Açık uçlu sorularda kullanılan kelimeler analiz edildiğinde ise ayırt edici özellerinden *emoji kullanımı* ile *cevapların miktarı ile sıklığı* ön plana çıkmaktadır. Tasarımcılar dışa dönük bir sohbet botu tasarlarken emoji kullanımı ve cevap miktarı ile sıklığı parametrelerini göz önünde bulundurmalılardır. Bunun yanı sıra, *konuşma şekli* parametresi de sohbet botlarının tasarlanmasında önemli bir kriterdir. Tasarlanacak olan kişiliğe özgün olarak botun konuşma şekli oluşturulmalıdır. Ayrıca, *konuşma isteği* parametresi dışa dönüklülük özelinde oluşturulmuş bir parametre olup tasarımcıların tasarlayacakları sohbet botunun kişiliği özelinde yeni parametreler oluşturmaya da açık olmaları gerekmektedir.

Bunun yanı sıra, farklı kişilikteki sohbet botlarının yetkinlikleri kişiliklerinden bağımsız olarak değerlendirilmiş ve aynı derecede yetkin görülmüşlerdir. Bu da 2. Hipotezi doğrulamaktadır. Bu sebeple, sohbet botlarının tasarlanmış kişiliklerinin yetkinliklerini etkilemediğini göz önüne alarak kullanıcılar oluşturacakları sohbet botlarında dışa ve içe dönük karakter özelliklerini kullanabilirler.

Katılımcıların kişiliklerinin bu sonuçlarda bir etkisi olup olmadığına bakıldığında bir etkisinin olmadığı görülmüştür. Bir başka deyişle, sohbet botları benzer veya zıt kişilikteki katılımcılar tarafından farklı algılanmamıştır. Örneğin, içe dönük bir katılımcı dışa dönük sohbet botunu dışa dönük katılımcılara göre daha dışa dönük algılamamıştır. Aynı şekilde, katılımcı kişiliklerinin *eğlence, kabiliyet, etkileşimin yararı* ve *sevme* boyutları üzerinde de bir etkisi olmadığı sonucu elde edilmiştir. Bir başka deyişle, katılımcılar iki farklı kişilikteki sohbet botunu aynı derecede eğlenceli, kabiliyetli bulmuş, kullanıcı ile olan etkileşimlerini aynı derecede yararlı değerlendirmiş ve ikisini de aynı derecede sevmiştir. Bu durumda, 3. Hipotez reddedilmiştir. Bunun nedeni olarak sohbet botlarının yeni gelişen bir teknoloji olması ve katılımcıların %70.1'inin günlük hayatında bu sohbet botların yer etmemiş olmaması söylenebilir. Birçok kullanıcının günümüzde sohbet botlarından beklentisi, istedikleri işlemi başarıyla yerine getirebilmesi olduğundan kişilik özellikleri şu an için kullanıcılar tarafından bir tercih olarak görülmemektedir denilebilir.

Bu çalışmanın kısıtı olarak katılımcılar ile çevrimiçi bir anket çalışmasının yürütülmesi görülebilir. Ne kadar çevrimiçi araştırma sınırlı bir süre içinde daha fazla katılımcıya ulaşılmasına yardımcı olsa da, katılımcıların tasarlanan sohbet botlarını ile bizzat etkileşime girerek deneyimlemeleri ve derinlemesine görüşmelerle verilerin toplaması araştırmanın sonuçları için daha faydalı olabilir. Bu nedenle, gelecekte bu tür bir çalışma yürütülebilir. Buna ek olarak, farklı zıt kişilikteki sohbet botları araştırılabilir. Bu çalışmada dışa dönük ve içe dönük kişilik özellikleri incelenmiştir; ancak, aynı metodoloji farklı zıt kişiliklerdeki sohbet botları için uygulanabilir. Örneğin, dominant-çekinik, resmi-gayri resmi, neşeli-melankolik kişilikteki sohbet botlarının araştırılması alana katkı sağlayacaktır. Ayrıca, sohbetler yeni gelişen bir teknoloji olarak kabul edildiğinden, insanların yaşamlarında henüz hayati bir yer bulamadılar. Bu nedenle, aynı çalışma gelecekte bir zamanda gerçekleştirildiğinde farklı sonuçlar elde edilebilir. Bunu gelecekte tekrar denemek ve bugünle karşılaştırmak değerli çıktı olacaktır.

Özet olarak, tezin amacı, kullanıcı ile duygusal bir bağ oluşturmak amacıyla metin tabanlı kullanıcı ara yüzlerinde, bir başka deyişle sohbet botlarında, kişiliğin önemini incelemek ve tasarımcıların bu konuda nasıl bir yol izlemesi gerektiğine yardımcı olmaktır. Sohbet botlarında kişiliklerin araştırılması tasarımcılara ve araştırmacılara yol gösterici olacaktır. Bu nedenle, potansiyeli olan bu konu hakkında daha fazla çalışmaya yer verilmesi gerekmektedir.

1. INTRODUCTION

1.1 Motivation

Personality in products has been investigated for many years on the view of product design, human-computer interaction (HCI), interaction design, communication, and marketing professions. Even if personalities are not deliberately applied to products, it can be said that users may give meaning to products according to their interactions. In the thesis, product personalities will be examined not only on the perspective of product design but also on the interaction design perspective. To be more specific, product personalities on conversational user interfaces -especially on chatbots- will be the focus of the thesis.

Today, with technological improvements, products have become more interactive. In daily life, it is not a surprise confronting a product that talks to you or texting you, which is called lifelike products. Moreover, it is shown that lifelike behavior in products builds a deep relationship between user and product. Thus, investigation about lifelike products should be considered by design researchers as a valuable research area for better user experiences.

Most of the lifelike products have a place in daily life as virtual personal assistants. Virtual personal assistants can be defined as assistants who help users for their daily activities. Siri, Google Assistant, Alexa, Cortana can be considered as virtual assistants who are worldwide famous. Siri is the first marketing-oriented virtual assistant who has been able to touch users' daily life (Apple, 2019). Moreover, these virtual personal assistants are examined under the name of "Conversational User Interfaces" (CUIs). In addition to that, CUIs are divided into two sub-categories, which are text or voice-based interfaces. Voice-based CUIs are named as voice user interfaces (VUIs) while text-based user interfaces are defined as chatbots. Even though chatbots have a terrible reputation as being less intelligent assistants than VUIs; in fact, the main difference between them is their dialogue systems which can be text-based or voice-based, not the technology they used in their infrastructure (Imrie and Bednar, 2013; Klopfenstein et al., 2017). Therefore, in the thesis, the name of the chatbot will be used as a text-based VPA.

Chatbots have several benefits for users and companies. For users, chatbots are agents that they can reach any time of a day on their devices. Chatbots are not only easy to use but also provide one-to-one communication. Also, users save money and time by using the chatbot service anywhere they want and get personalized offers by chatbots. For companies, they provide direct customer service to their customers anytime and also can collect personal data of their customers such as what kind of problems customers face and pain points of their business for specific customer segments. Their communication method becomes not only automized but also personalized with decreased costs for customer services (Zumstein and Hundertmark, 2017). Therefore, an increase in the number of chatbots will be inevitable.

Reports show that the number of chatbot in the market has been rising day by day. The contribution of Facebook Messenger cannot be denied. As Zumstein and Hundertmark (2017) stated, the number of chatbot in Facebook Messenger was 11 thousand in June 2016, and on April 2017, this number reached 100 thousand (Facebook, 2017). Besides, the potential revenue of chatbot transactions has significantly increased to annual firms is also declared very high (Business Insider, 2017). Therefore, chatbots should be considered as an essential research topic for designers in order to provide better user experiences.

Personality in chatbots is also a potential area for research because nowadays, people confront chatbots in their daily life and, deliberately or not, behaviors of chatbots can interpret as a characteristic feature of a personality. Many research indicates that personality in conversational user interfaces is a critical issue in order to establish a sense of trust on users (Lee and See, 2004). It is asserted that the establishment of a trust-based relationship can be possible with the use of personality in CUIs, and it strengthens the communication between the user and the product. Also, it can be said that trust contributes to the construction of the bond between user and product. This bond also can create user engagement for the product and brand.

Another reason investigating chatbot personalities is to provide strengthening ways of user engagement with the emotional bond between user and product. With the development in the technology, perception and requirement of future products have changed. Norman (2004) mentions that emotional attachment is a significant factor for designing future products in his book. Giving personality to conversational user interfaces can be a method for the creation of emotional attachment between product and user. Therefore, it is believed that investigations about this topic will contribute to the design of products with CUI significantly.

Since chatbots popularity has been increasing only for the last years, there is not a significant amount of study on chatbot personalities yet. Zumstein and Hundertmark (2017), for instance, studied on chatbot usage on the public transportation sector. McTear, Callejas and Griol (2016) investigated conversational user interfaces in multiple perspectives including personality; but, their study is more like a theoretical guide for designer and developers. In addition to that, Smestad's study (2018) aims to show that giving personality traits to chatbots improves user experience in the right way. However, the impact of chatbot personalities on different user personalities has not been studied in the case of chatbots. Therefore, personality studies and similarity attraction in chatbots should be conducted in order to present more concrete and specified outputs for designers.

Generally, the reason why chatbot personalities should be considered as a valuable research area is that creating better experiences on the conversation interfaces is very important for user satisfaction and this is a critical issue for both users and companies. Therefore, as the role of a designer is establishing a better relationship between user and product, chatbots should be considered as a novel and valuable research area for design. In the thesis, personality in chatbots examined to increase the quality of human-computer interaction. For this purpose, how can product personality be built on chatbots for target users will be discussed in the thesis.

1.2 Similarity Attraction

In psychology, similarity attraction is a term to describe that people tend to attract people who similar to them. Similarity-attraction paradigm is defined by Bryne (1971) and, then, studies about similarity attraction have been conducted through a variety of different sub-similarity factors such as attitude, value, construct, structural, and personality for deep relations (Neimeyer and Mitchell, 1988). Personality similarity on relations has been studied not only in psychology but also in HCI and HRI. In the thesis, also, personality similarity will be the main focus of the investigation.

Bryne, Griffitt, and Stefaniak (1967) stated that there are various studies and opinions on this subject. For instance, Izard (1960) advocate the attraction of similarity; in contrast, Rychlak (1965) advocate the complementarity attraction. In addition, Secord and Backman (1964) supported the combination of the two, and Rosenfeld and Jackson (1965) partially supported similarity attraction. Apart from these, Hoffman (1958) did not think that there is a relationship between similarity and attractiveness. However, most of the personality studies show that there is a connection between personality and attraction level of users. However, in the case of chatbots, there are no investigations about similarity attraction. In the thesis, to fulfill the gap in the field, a user study will be conducted on chatbot personality, and similarity attraction on chatbots will be examined.

1.3 Research Question

Researchers have investigating giving specified personality characteristics on products such as robots and websites, and users' perception of these personalities. Even personality in chatbots has been examined, applying specific personality type on chatbots and examining users' reactions is missing. Therefore, extroverted and introverted personality types will be applied to chatbots, and user perceptions about these personalities will be examined.

Hypothesis and research questions are listed as:

Participants will be able to recognize chatbot personalities (H1).

RQ1: Do participants consider the extroverted chatbot as an extrovert, the introverted chatbot as an introvert?

RQ2: What are the design criteria for extroverted and introverted chatbots?

Personalities of chatbots will not affect the perception of chatbots' capability (H2).

RQ3: Is there a significant difference in *competence* and *usefulness of interaction* dimensions between the chatbots?

The personalities of participants will affect how do they perceive chatbots with different personalities. (H3).

RQ4: Do the personalities of the participants affect the perception of chatbot personality?

RQ5: Do the personalities of the participants affect their perception of *fun, competence, the usefulness of interaction*, and *liking* dimensions of the chatbots?

First research question (RQ1) aims to measure whether the right path was chosen for designing personality in the chatbots. Similarity attractions studies also have this kind of measurement questions on their investigations. However, extroversion and introversion personalities on chatbots have not been examined before. Moreover, the results will indicate whether the parameters used in chatbots is useful to create a personality, and these parameters can guide designers to create a personality in a CUI. A field study was held for these two and subsequent research questions. On the method chapter, the scenario and the procedure that conducted will be examined.

In the second research question (RQ2), application methods for extroverted and introverted characteristics were discussed in the conversational user interface perspective. For this aim, literature was reviewed, and examples were discussed. In the literature, there is a personality study in particular for chatbots (Smestad, 2018). In this study, one of the chatbots was attempted to give personality. Even if there is no intentionally designed personality in a product, it is always perceived as having a personality by users (Dryer, 1999). For this reason, it is vital to determine design criteria for specific personalities and apply over products in personality research. Therefore, in the thesis, design criteria to create a chatbot personality was focused. The design criteria mainly examined by extroverted and introverted chatbots. Verbal and non-verbal design criteria have been established within the scope of the media.

Even the chatbots in the research have different personalities, their capability for their duties are kept the same. Whether their level of capability was perceived similar by participants, the third research question (RQ3) aims to measure the perception of *competence* and *usefulness of interaction* dimensions of the chatbots. Nass et al. (1995) also had a similar goal on their research, but their investigation was about dominant and submissive text-based interfaces. Although Lee and Nass (2003) studied on similarity attraction on extroverted and introverted interfaces, their investigation was about voice-based. Therefore, in the RQ3, capability perceptions on chatbots which have the contrast personalities in mainly will be researched.

There are various studies about product personality not only in HCI (Nass et al., 1995; Lee and Nass, 2003), HRI (Aly and Tapus, 2015; Tapus, Tapus and Mataric, 2013) but also specifically in voice user interfaces (Mennicken, 2016), chatbots (Thies et al., 2017; Smestad, 2018). While extroverted and introverted personalities on interfaces or social robots have been examined; however, none of these studies has worked on similarity attraction on chatbots, which can increase emotional bonds and trust in users. Therefore, similarity attraction on chatbots can be considered as a gap in the literature. Following research questions of the thesis (RQ4 and RQ5) focuses on fill this gap in the literature. In the fourth research question (RQ4), it was aimed to examine whether the personalities of the users have an impact on the perception of the chatbot personality. In that way, it can be measured whether the designed personalities on chatbots are universal or vary from user to user. Also, in the final research question (RQ5), whether users' preferences depend on their personalities will be measured and discussed. For this aim, perception of *satisfaction, liking*, *competence*, and *usefulness of interaction* dimensions between the chatbots will be compared, respect to personalities of participants.

2. LITERATURE

2.1 Lifelike Products

The term of lifelike behaviors has been coined at the literature in the last years. Although “lifelike behaviors in products” the word pattern (Kestner, J. et al., 2009; Togler, J., Hemmert, F., and Wettach, R., 2009) have been mentioned literally on their papers for the first time, research about living products (or interfaces) have been discussed with different words in different contexts for many years. Human-robot interaction (HRI), human-computer interaction (HCI), interaction design, product design professions use the various term to describe lifelike products that act like alive organisms. Notably, social robots have an impact on the creation of the research area. Investigations in HRI have challenged developments in human-like and pet-like robots. Besides HRI, HCI and product design have also contributed to this topic on different perspectives. Conversational user interfaces can be accepted as one of the most lifelike products users encounter in daily life.

2.1.1 Conversational user interfaces

In recent years, there have been emerged many conversational user interfaces in order to serve users in particular topics or fields in everyday life. Even though different names are used for describing conversational user interfaces such as smart assistants, conversational agents, digital assistants in the literature; conversational user interfaces will be used as a term in the thesis. Conversational user interfaces are assistants that communicate with the user through voice or text-based conversations.

McTear, Callejas and Griol (2016) examine conversational user interfaces under 4 group, which are virtual personal assistant, social robot, embodied conversational agents, and chatbots. Social robots are mentioned as human-like or pet-like robots that interact with people by mimicking social behaviors and following social rules. They can serve in many parts of daily life like entertainment, companionship, or education. Embodied conversational agents (ECAs), on the other hand, are defined as “computer-generated animated characters”, and they simulate human facial expressions, body movement, and speech. ECAs are generally used for transmitting informative contents in fields like healthcare, education since they are perceived as trust-worthy characters. McTear, Callejas and Griol (2016) also describe virtual personal assistants (VPAs) as agents that communicate users in order to help them for their daily activities like finding information that user needs, editing calendar, booking flights, setting the alarm and entertaining users with chit-chat. Chatbots are declared as bots that perform casual conversations instead of completing specific tasks as an automated online assistant.

Moreover, McTear, Callejas and Griol (2016) assert that chatbots are small talk oriented and generally have simple infrastructure systems. There is a misdirection on their study that VPAs are more intelligent than chatbots; however, the intelligence of both agent types depends on their process systems, techniques, and their databases. In other words, VPAs and chatbots can have the same information technology (IT) infrastructure depends on companies' business decisions, and that does not mean chatbots are less intelligent than VPAs. Chatbots can be defined as a text-based subcategory of virtual personal assistants and will be examined in this perspective in the thesis (Figure 1.1).

Moreover, conversational user interfaces can be examined in 3 categories by its communication tool; which are text-based, voice-based, and text and voice-based. For example, Siri, Apple's voice-based VPA, communicates with users via voice and users also should speak when asking a question to Siri. Google Assistant, moreover, offers two options for smartphone users: text and voice. It means that Google Assistant's users can prefer to communicate via voice or text. Chatbots, on the other hand, can be described as a text-based VPA and they generally live on a specific channel like Facebook Messenger, WhatsApp, or Slack. Therefore, the general VPA concept includes both voice and text-based personal assistants, while chatbots are considered as a text-based specialized VPA (Imrie and Bednar, 2013; Klopfenstein et al., 2017). In the thesis, chatbot the term will be used to describe text-based virtual personal assistants. To be more explicit, first, virtual personal assistants, in general, will be examined; then, chatbots, mainly, will be the main topic of the discussion.

2.1.1.1 Virtual personal assistants

Virtual personal assistants can be stated in several names such as Conversational Agents, Intelligent Personal Assistants, Personal Digital Assistants in the literature. As mentioned above, VPAs are assistant that help to manage users' life by accomplish tasks or giving information by using voice or text-based inputs (Baber, 2002). Today, many technology companies have been developing their virtual personal assistants to serve their customers. Apple's Siri, Google Assistant, Amazon's Alexa can be examples for few and most popular of them. Also, it is predicted that the market for VPAs will be raised to $4.61 Billion by the 2020s (Kamitis, 2016). Therefore, VPAs should be considered a critical topic for investigations.

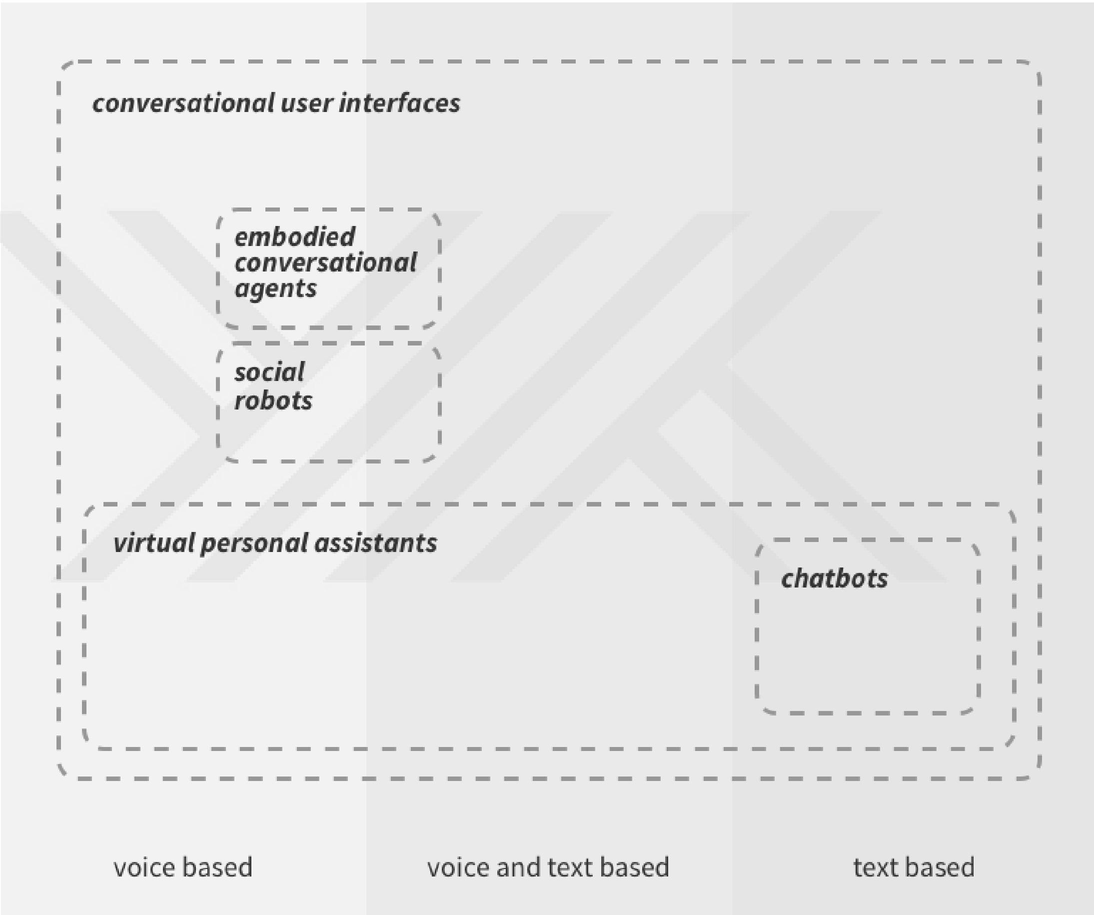

Figure 2.1 : Subcategories of conversational user interfaces

Siri, in 2011, is the first released VPA in the market, and Apple (2011) introduced Siri as "an intelligent assistant that helps you get things done just by asking". The difference between Siri and other AI products in the market at that time was that Siri did not just accomplish her tasks; she also communicates with users, based on social rules. (Guzman, 2017). So, users can interact with their VPA Siri via their iOS software systems whenever they want to through their devices. Moreover, smartphones are not the only device that VPAs can be communicated; they are also in our homes or cars. For instance, Google Assistant is also a VPA which is developed for assisting users' demands like calling friends, ordering pizza and can be activated whenever by saying "OK Google" to your android. Google Assistant can also be connected through Google Home, which is a device that specialized for home use. As a result, it can be interpreted that life-like products have become more into our daily lives, not only in our smartphones but also our home.

2.1.1.2 Chatbots

A chatbot, also known as a bot, chatterbot or conversational agent, is a text-based artificial intelligence which receives users' textual input, understand their intents and returns outputs for these specific intents depends on rules and data. Chatbots are a member of the virtual personal assistant family because they can be deployed the same natural language understanding (NLU), which can be accepted as the brain of VPAs, in their infrastructure as other VPAs. Although it is argued that chatbots are bots that accomplish rule-based autonomous small talks and not assisting, in fact, it is revealed that chatbots can perform tasks that facilitate users' lives, can assist them as much as voice user interfaces, and the only difference is that they use written communication tools (Zumstein, 2017). Therefore, chatbots should be accepted as text-based VPAs. In fact, it can be said that chatbots are the first version of virtual personal assistants.

The beginning of chatbots is a bot called ELIZA, which is produced by a therapist called Weizenbaum (1966). ELIZA is designed to help people with their psychological problems like a psychologist in a conversation with a humanoid language and will set an example for subsequent chatbots. Today, chatbots serve in many areas. Statistics show that real estate, travel, education, healthcare, and finance are the top sectors benefiting from chatbots (Chatbot 2019 Trends and Stats with Insider Reports, 2019). These VPAs aim to provide information that users need in a human-like dialogue. The primary motivation of the firms, in this regard, is to provide such assistance to replace the call centers and provide an automated standard service to the users (McTear et al., 2016).

Chatbots generally interact with their users through channels like Facebook Messenger, WhatsApp, Slack, or websites. For example, BabyCenter UK's chatbot on Facebook Messenger helps users who have concerns about their baby. According to complaint type and age of the baby, the chatbot offers suggestions to parents. Also, the chatbot sends information to the users periodically according to the growth process of their baby. Another Facebook Messenger example can be 1800Flowers, which is a chatbot designed to manage flower orders. The service process proceeds as follows: first, the chatbot asks for an address for the delivery; after the address confirmation, the user picks a type of flowers which suggested; finally, the users decide arrangements of the flowers from looking carrousel cards and the order is ready for delivery. Most of the dialogs occur via buttons or carrousels in order to suggest options to users and decrease the number of mistakes.

Recently, WhatsApp Business decided to allow companies to provide services for their customers on the channel. An example of WhatsApp chatbot can be KLM, Royal Dutch Airlines (KLM, 2019). KLM is requesting permission from their customers to communicate with them via WhatsApp. The KLM's main aim in creating this chatbot service is to support their customers during their travel time. First of all, a copy of the reservation is sent to the users completing the reservation via the WhatsApp along with the travel details. The chatbot, then, sends a reminder to the passenger to check-in. After check-in, a digital boarding pass is forwarded to the customer. The chatbot also transmits status to the passenger via WhatsApp when there is an update on the user's flight. Likewise, any request from the passenger or a question can be contacted here with KLM. As a result, it can be said that KLM has digitized the entire process of its users through the chatbot.

Today, in 2019, 5 billion of these users use messaging applications (GSMA Intelligence, 2019) and there are also over 4.3 billion internet users worldwide (Digital 2019: Global Digital Overview, 2019) This data can be accepted as an essential indicator for the potential of chatbots in the market. In addition to that, the number of chatbots has been already increasing year by year. For instance, the number of chatbots on Facebook raised to 300.000 (Johnson, 2019). That is why, chatbots should be investigated for the sake of users' since it will be not unavoidable to contact with digital assistants in the future. For better interactions, this topic has a significant issue for design researchers.

2.2 Personality Theories

Variety of personality theories makes it compelling to determine the right personality method for the research. Some of the theories say the same things in different words, and others say different things in the same words (John and Srivastava 1999). Thus, analyzing theories and measurement methods for personality investigations is a critical issue in order to be accurate on the research results.

There are several taxonomies for personalities based on psychological theories such as Eynseck Personality Theory (Eysenck, 1946), Kiesler's Interpersonal Circle (Kiesler, 1982). On the other hand, Cattell's 16 Personality Factors (Catell, 1943), The Big Five Model (Goldberg, 1990) is not derived from theoretical context, but people's description about themselves were used for shaping big five taxonomies (John and Srivastava, 1999).

Eysenck Personality Theory (1991) declared that personalities could be had interpersonal dimensions, which are called Introversion/Extroversion and Neuroticism/Stability. Eysenck also determined personality traits according to the two axes of extroversion and neuroticism. Neuroticism is the interpersonal dimension of neurotic stability, while extroversion is the interpersonal dimension of expressiveness. The intersection areas of the two axes tell us the type of personality of the person and its traits. For example, the intersection of extroversion-stable means "sanguine" personality type, which contains traits like sociable, talkative, easygoing, and leadership. Unstable side of extroversion axis, which is named as "choleric", has traits such as active, changeable, aggressive, and touchy. On the other hand, introversion side of the coordinates is separated as "phlegmatic" (introvert-stable) and "melancholic" (introvert-unstable). Phlegmatic personality is described with traits like calm, reliable, thoughtful, and passive, while melancholic personality type traits are shown as quiet, pessimistic, rigid, and moody (Eynseck, 1991).

On the other hand, Kiesler's Interpersonal Circle (1983) is created for personality research and supports an interpersonal structure like Eynseck. Kiesler (1983) asserts that people interact with each other over two main criteria; control and affiliation. Therefore, his circle consists of these two-main axes. The control axis can also be named as dominance, and the close relationship axis can be called friendliness. Moreover, as Mahalik (2000) states, Kiesler's Interpersonal Circle contains three nested circles. The outer circle shows extreme personality behaviors, the moderate personality behaviors in the middle, and the mildest behaviors in the inner ones. Every circle has 16 personality traits. For example, traits of the inner circle are named as follows; dominant, assured, exhibitionistic, sociable, friendly, warm, trusting, deferent, submissive, unassured, inhibited, detached, hostile, cold, mistrusting, and competitive (Kiesler, 1983).

Personality taxonomies of both Eynseck’s Personality Theory and Kiesler’s Interpersonal Circle based on theoretical frameworks; however, these taxonomies seem abstract for personality investigations. Therefore, in order to conduct more systematic personality research, it is necessary to look at lexical taxonomy studies (John and Srivasta, 1999).

2.2.1 Catell’s 16 personality factors

There are several researchers for the construction of personality factors by using natural language as a tool instead of a theoretical framework. For example, Klages (1926), Baumgarten (1933), Allport and Odbert (1936) studied on personality dictionary.

Allport and Odbert (1936) first coined the idea of a taxonomy of traits for psychological personality types. Allport and Odbert, first, selected18.000 terms from the dictionary and grouped them under four headings. The terms that are chosen consist of words used to describe a human, and they are grouped under terms of “traits”, “states”, “activities”, and “evaluation”. Traits defined as terms used to describe consistent and stable modes of adaptation to the individual’s environment. States can be explained as a temporary mood of individuals. Activities consist of the words used in evaluating people like average, excellent. Evaluations include the terms in the definitions made over the physical characteristics and abilities of individuals.

Then, Norman (1967) try to a similar method to conceptualize personality types into restrictive subgroups, which contains total 1710 traits in seven categories which are “traits”, “internal states”, “activities”, “physical states”, “effects”, “roles”, and “social evaluations”. However, neither Allport and Odbert’s nor Norman’s categories were not seemed as eligible for applying on personality research since the amount of the traits were so much for measurement (Allen & Potkay, 1981) as stated by John and Srivastava (1999). Then, Chaplin et al. (1988) separated groups as “traits”, “states” and “activities”. It is described as traits are permanent where states are temporary moods of personality.

Moreover, Cattell (1943) reviewed Allport and Odbert’s 4500 traits and classified it into 35 variables. John and Srivastava (1999) state that after factor analyses, first, 12 personal factors emerged; then, eventually, it became 16 personal factors (Cattell, Eber, and Tatsvoka, 1970). However, factors were criticized by another researcher for being inaccurate (John and Srivastava, 1999). For example, Digman and Takemoto-Chock (1981) state that "original model, based on the unfortunate clerical errors noted here, cannot have been correct". After the adjustment, a new version of 16 Personality Factors was created, which inspired by The Big Five perspective.

2.2.2 The big five

The Big Five, also known as The Five Factor Model, is believed that the most common version of personality taxonomy in the literature. First, Fiske (1949) developed a taxonomy that is very similar to The Big Five through factor analyzing Cattell's 22 variables. Tupes and Christal (1961) re-calculated correlations on samples, and they concluded that there are five influential factors. Many researchers replicated the Five Factor taxonomy (John and Srivastava, 1999).

The model's last version developed by Goldberg (1990) by using Norman's (1967) 1710 trait list under 75 linguistic categories, in order to build more comprehensive personality factors. After factor analyzing, Goldberg represented the first five factors as Big Five. Various research was also conducted to show the validity of these five factors (Goldberg, 1990; Saucier and Goldberg, 1996b; Saucier, 1994).

The five factors can be listed as Extroversion, Agreeableness, Conscientiousness, Neuroticism, and Openness to Experience and explained below (Novikova, 2013).

Extroversion: The factor can also be called Surgency. While personalities high in extroversion are more talkative, excited, energetic, expressing their emotions comfortably and assertive personalities; introverted people are defined as more reserved, silent, expressing their emotions uncomfortably and submissive.

Agreeableness: Adjectives like kind, sympathetic, warm, cooperative, compassionate are related to people who have a high score on agreeableness factor on The Big Five Model. Contrast to a high score on this factor, personalities which are disagreeable more tend to be more competitive, skeptical, and opponent.

Conscientiousness: People who score higher the conscientiousness factor are more prone to being disciplined, tidy, planned, and trying to fulfill their goals in life. Those who have a low score are more careless in their plans and have less self-disciplined personalities.

Neuroticism: The factor name of Emotional Stability is sometimes used for this factor. In general, personalities with high neuroticism have more sensitive emotions, tend to be easily upset, worry, and they are vulnerable to stress while the opposite personalities are more likely to be calm and self-confident against to stressful situations.

Openness to Experience: It can be also known as Intellect, Openness, or Culture. Those who have a high score on Openness to Experience factor tend to be more creative, open to new ideas and adventures, and experience their emotions intensely. Contrast to that, those who have a low score on the factor are more conservative for new ideas or adventures and have a poor emotional way of life.

So many researchers created personality models similar to the Big Five (e.g., Botwin & D. M. Buss, 1989; Digman and Inouye, 1986; Field & Millsap, 1991; John, 1990; McCrae and Costa, 1985a, 1987; Peabody and Goldberg, 1989). For many years, researchers seek a personality guideline to analyze their data. In the paper, called "Paradigm Shift to The Integrative Big Five Trait Taxonomy", method choices of researchers were listed and compared as Cattell/Eysenck or Big Five/FFM in order to show that which personality taxonomy more preferable on investigations. The Big Five personality trait publications have a dramatic increase than Cattell/Eysenck (John, Naumann and Soto, 2008). Moreover, it is stated that Big Five Taxonomy collects all taxonomies under one roof (John, 1990). Thus, on the research, Big Five Taxonomy will be considered as a guideline.

2.3 Personality in conversational user interfaces

The use of personality in conversational user interfaces has been also discussed in different fields. For example, some research investigates conversational user interfaces for educational purposes (Bevacqua et al., 2010; Tewari, Liu, Cai and Canny, 2012) and some of them for elder or mentally disabled people (Yaghoubzadeh, Kramer, Pitsch and Kopp, 2013). In mainly, personality studies (e.g., Bates, 1994; Hayes-Roth et al., 1995; Maes, 1995; Nass et al., 1995b; Oren et al., 1990) have been emerged to enable conversational user interfaces to have a better social interaction with users, as

Dryer (1999) states. Interactions in these studies were measured by giving specific emotions, personalities, or roles to the residents.

Anthropomorphism, which can be described as mimicking human forms and behaviors, has an immense role in the creation of these personalities and user engagement (Dautenhahn et al. 2002, Epley et al. 2007, Lee and See 2004). Most of the conversational user interfaces benefit from the anthropomorphic approach in order to create human-like interfaces. Moreover, this anthropomorphic approach is essential not only for the visual elements of the product but also for its behavioral elements. For this reason, it can be asserted that personality is one of the main points of this anthropomorphic approach because the product personality forms the basis of these behavioral characteristics (Lee and Moray, 1992). Both verbal and non-verbal parameters are used to give interfaces human-like personality characteristics. Therefore, it can be asserted that personality characteristics help to create these anthropomorphic CUIs.

Culley and Madhavan (2013) assert that trust is a very critical point for human-computer interaction due to the fact that users' satisfaction determines the performance level of the service (Bartneck et al., 2009). Users perceive the competence and capacity of the services from designed interfaces, and their judgment about the service depends on trust between them. In other words, no matter how advanced a product is and how competent to meet the needs of the user, if it is unable to establish a sense of trust and cannot establish a bond with users; it will not be able to create the perception of that is intended. System capacity and performance may be interpreted as lower than the actual. Therefore, this situation should be given importance in order to establish deep relations between products and user. Moreover, using anthropomorphism on interfaces builds a sense of trust between the user and these interfaces by creating similar characteristics with users (Lee and See, 2004). As a result, it can be noted that anthropomorphism is used for the creation of CUI personalities in order to build deep relations with users.

For example, most of the conversational user interfaces have a name and avatar. Also, some of them are designed according to specific gender and age. These elements are used for the creation of personality and affect users' perception of their personalities. For example, Siri has a female voice by default, and her voice can give us a clue for understanding her age scale. These kinds of clues affect perceptions of personalities because western societies are shaped by socially constructed bias for issues like gender. For instance, studies (Nass and Moon, 2000) show that users think that male-dominant computer voices are friendlier to them than female ones. In other words, judgmental responses from female voices are perceived as bossier.

Moreover, the level of intellectuality of computers was also perceived differently by topics and gender. For example, although the information provided by two types of computers was the same, the female voice computer was considered more knowledgeable about love and relationships which are stereotyped subjects for women. It shows that even if the intellectual level of artificial intelligence is the same; their identities, naturally their personalities, play an essential role in user-product engagements. Therefore, the creation of the personality in CUIs should be designed according to the target group.

All sample products that are examined in the thesis have the same concern, which is to create better engagements with users. So, verbal and non-verbal communication tools have been used in order to give the feature of human-like characteristics to these products. Verbal communication tools can be described as text or voice usage in product personalities. Gestures, posture, expressions, visual appearance, and melodies are discussed under the non-verbal communication tools in the literature. In the thesis, personality in text-based conversational user interfaces will be examined under not only verbal tools such as tone of voice, verbal expressions but also some of the non-verbal tools such as visual appearance, amount of answers and frequency, and visual expressions.

3. RELATED WORKS

3.1 Product Personality

Researchers in the vary of perspectives have investigated the creation and perception of product personality. Gover (2006) and Jordan (2002) studied product personalities in the perspective of visual appearance. In the physical interaction context, Desmet et al. (2008) investigated product personalities on dynamic interactions through examining personality traits as dominant and elegant. The result of the study shows that traits can be given to products through five aspects of physical interaction (Jenlert and Stoltman, 1997) which are force, sound, motion, texture, and performance and results of the experiment explain that dominant and elegant product personalities can be created by interaction.

In addition to that, Lee et al. (2006) researched personality in robots which are extrovert and introvert and analyzed tangible interaction characteristics. Verbal (voice, rate of speech) and nonverbal tools (gestures, movements, facial expressions) were used for the creation of personality in robots, and whether personalities were perceived as intended was measured. Results show that personality characteristics can be applied to robots through verbal and nonverbal cues. In the perspective of HCI, Nass et al. (1995) proceed an investigation through text-based feedbacks from computers which have a personality like a dominant or submissive one. The study asserts that little changes in verbal communication affect personality and computers can have a personality like humans.

In a study conducted within the context of personal virtual assistants, it was examined how Siri's personality was perceived by users (Cowan et al., 2017). In general, the participants found Siri "sassy" and "friendly", and some participants stated that the personality of the VPA influenced their relationship with them and did not want to upset Siri. Besides, Siri's personality was evaluated in the cultural context. The participants shared their experiences and thoughts by talking to Siri in different accents. For example, the Japanese Siri was seen as more serious, polite, and less prone to joke (Cowan et al., 2017). It means that changes in the cultural context also affected the perception of the personality.

The reason for the investigations about product personality is to create deep relations between users and products and present to users better experiences, as mentioned before. For this aim, many fields have been conducted studies and not only verbal but also non-verbal communication tools on products have been examined. Also in the thesis, besides verbal, nonverbal tools will be examined in the case of conversational user interfaces.

3.2 Similarity Attraction

There are various investigations about product personality that provide services like robots, web sites. These investigations focus on the perception and effects of product personalities on users. Moreover, the controversial hypothesis has been examined in the field of human-robot Interaction (HRI) and human-computer interaction (HCI), which is named similarity attraction (Byrne et al., 1986; Isbister & Nass, 2000). Similarity attraction can be explained as that users are more inclined to like better a product that has a personality similar to themselves.

For instance, Nass et al. (1995) examine user preferences on computers through testing dominant and submissive personality types. In the test, participants are assigned a task and complete this task by getting help from the computer they randomly match. One of the computers is dominant, and the other is recessive. After completing the task, the participants are asked about the computer they are paired with, and as they complete the tasks, it is learned how much the computer they interacted with help them and how much they liked it. The results show that the dominant participants prefer to use the dominant computer more than those with submissive personalities. Similarly, those who are submissive also have similar results.

Moreover, a similar study was conducted on a shopping assistant case (Al-Natour et al., (2006). This study, as in the previous study, investigated dominant and submissive characters and similar results were obtained. As a result, it can be said that users prefer to get help from a service that matches their personalities.

Another study also shows similar results; for example, in this study of Lee and Nass (2003) relationship with the personality types of the sound has been examined through web sites. Two experiments were carried out in this research on extroverted and introverted personality types. In the first of the experiments, the book descriptions on a site selling books were voiced by two different personality types, and the manner of saying was changed according to personalities by keeping the words and sentences fixed. In the second experiment, the product descriptions are given in a similar site with the auction. In this experiment, the words used were changed, and the manner of saying kept still. The users were asked to perform the scenarios by listening to these voice recordings in both experiments. As a result of the two experiments, it was observed that the participants with the extroverted personality were better at distinguishing personality types and preferred the sound close to their personalities. Likewise, in the results, introverted participants prefer the sound that is close to their personality compared to the opposing personality. As a consequence, product personality can be created by multiple tools such as text, voice, manner of saying.

Besides HCI, HRI field also has investigations about similarity attraction on robot personalities through verbal or nonverbal communication tools. One example can be the research conducted by Aly and Tapus (2015). In this research which investigates the place of the personality in the human-robot relationship, the task of the robots is giving information about the restaurants in New York. While a robot exhibits this extroverted personality, the other one behaves like introverted personality; and not only words but also gestures of the robots were designed for the experiment. As a result, the similarity attraction in the personality was demonstrated in this experiment, too.

The last example for similarity attraction study in HRI is about assistant robot for rehabilitation (Tapus, Tapus and Mataric, 2013). Four rehabilitation exercises were given to the participants. While completing these, the robot next to them is found to motivate them. A group of participants works with an introverted robot, where others work with an extroverted robot. Robots say motivating words to participants to complete tasks during the study. They do this in a way that differs according to their personalities. For example, the extroverted robot is designed to be a more compelling robot, while the introverted one has a more nurturing role. As a result, also in this study, participants prefer the robot who is close to their personalities. It can be summarized that even voice and words are enough to create personality in products and build better relations between user and product.

On the other hand, another hypothesis declares that users are more attracted to personalities that complementary to theirs, which is called complementary attraction (Byrne et al., 1986). There are some studies about this hypothesis, and examples show that participants tend to choose complementary personality. For instance, on a study (Isbister and Nass, 2000) apply a similar scenario with a previous study (Nass et al., 1995), this time with an expressive character, and beside the previous study following the similar attraction, in the results of this study has the opposite data. Participants are more attracted to complementary personality in this version of the study.

In another example, the study aims to learn the preferences of the participants by interacting with a robot named AIBO, who has extroverted and introverted versions (Lee et al., 2006). It is seen that the participants have more tendencies to the robot, which is the opposite of their personalities. Compared to the previous studies, the results of the investigations can be presented as the reason for the tasks of the assistants. Joosse et al. (2013) explained this with a view of stereotypical professions. In their example, they assume that participants expect an extroverted character from a tourist guide while an introverted character from a housekeeper. However, no significant data could be collected in this study. The reason for this can be explained by focusing on how the position of the assistants in the task rather than in what task is positioned. In other words, the role of the assistant in the task may affect the personality preferences of the participants. For example, it may be expected that the assistant will have a personality close to his/her personality as it is expected to be a second mind from an assistant who offers a guiding service. In a friendly interaction or dialogue, one may expect the assistant to have an opposite personality to avoid suppressing his or her personality.

Finally, it can be summarized that most of the assistant services aim to help users, therefore, whether online or physical assistant, should be designed personality-oriented and these personalities should be considered the perspective of similarity attraction. With this approach, deep relations with users can be established, and assistants can place the everyday life of users. Thus, in the thesis, similarity attraction in smart assistants will be the main focus.

4. METHOD

It was decided to conduct a test with users in order to verify the hypotheses and to answer the research questions. In this study, two chatbots were designed; one of them was an extrovert, and the other one was an introvert. The literature was used to create these personalities, and the differences were created with specific variables. Then, the hypothetical dialogs with a user were created and animated according to the designed scenario. Chatbots carried out the same tasks within the framework of different personalities.

The user study was conducted online. Participants participating in the study assessed their personalities first. The reason for this was to examine whether the personalities of the participants had an impact on the assessment. Afterward, the participants watched one of the two versions of the chatbots, which corresponded to them randomly, and then made an overview of the chatbot. Thus, it was measured whether the personalities and capabilities of the chatbots were understood as intended and what the differences were.

In this section; firstly, the tools used in the measurement and what is aimed to measure on the user study were discussed. Then, the variables that are effective in the design of chatbot personalities, the scenarios of the videos, the participants, and the procedure were explained.

4.1 Personality Measurements

To get more accurate results, first, personality measurement tools in psychology should be discussed. For this aim, main personality measurement tools which are developed in the perspective of The Big Five will be examined in this section of the thesis and which tool was decided to apply will be explained. Then, what was aimed to be measured, what are the main dimensions of the measurement, and how it was planned to be measured in the study will be examined in details.

NEO Personality Inventory (NEO PI), Costa and McCrae (1985) first created a questionnaire to measure personalities among factors of neuroticism, extroversion, and openness to experience. Then, their new revised questionnaire, which is called NEO PI-R, was more comprehensive and had 240 items with agreeableness and conscientiousness questions and every Big Five factor has six facets on their inventory (Costa and McCrae, 1992). For example, extroversion is subcategorized as Gregariousness, Assertiveness, Activity, Excitement Seeking, Positive Emotions, and Warmth. Moreover, due to the length of the last questionnaire, Costa and McCrae (1992) developed NEO-FFI that contains 60 items in order to measure personalities. However, the questions on the inventory contain sentences, and it is expected from participant to fulfill the test according to their behaviors as an individual. Asking personality questions through humanistic way like on the NEO-FFI can make participant misdirect on their decisions about chatbot characteristics. For example, "I am a party person" is one of the questions of the questionnaire, which is not an appropriate reference for determination of chatbot personalities.

Moreover, Goldberg (1981) mentioned that a daily life language has personality codes hide in its nature; so, adjectives of the language can be used for describing personality characteristics. Thus, the usage of personality traits can be an alternative way to determine personalities. For this aim, there are two type tests in order to measure personalities with using specific trait words; the Big Five Markers and the Mini Markers. Both derive from the same personality measurement approach which categorizing personality types under the Big Five understanding.

The Big Five Markers (100 Unipolar Trait Descriptive Adjectives, TDA), according to the Big Five, Goldberg (1992) developed trait scale which is called the Big Five Markers, also known as unipolar trait descriptive adjectives. It has 100 items for testing personalities among Big Five factors that are extroversion, agreeableness, openness, neuroticism, and conscientiousness. Every factor has 20 total items, ten positives and ten negatives for each and uses a 9-point Likert scale from extremely inaccurate to extremely accurate. For example, measurement markers under extroversion factor are extroverted, talkative, assertive, verbal, energetic, bold, active, daring, vigorous, and unrestrained. Reversed adjectives for Extroversion factor, which can be also named as Introversion factor, are introverted, shy, quiet, reserved, untalkative, inhibited, withdrawn, timid, bashful, and unadventurous.

Based on universal research such as Chinese (Yang and Bond, 1990), Czech (Hrebickova, 1995), Italian (De Raad, Di Blas, and Perugini, 1998), Russian (Shmelyov and Pokhil'ko, 1993), and Turkish (Somer and Goldberg, 1999); it can be said that the Big Five Model is a universal model for local use. Only the fifth factor can be affected by its region (John, 1999). On the research, however, there is no reason for not using Turkish adjectives to determine personalities according to Big Five since our main factors, that will be measured, is the extroversion factor.

Mini Markers, on the other hand, is an abbreviated version of TDA. It contains four positive and four negative adjectives for each factor. For examples, under the extroversion factor; there are adjectives such as bold, extroverted, talkative, energetic, bashful (reversed), quiet (reversed), shy (reversed), and withdrawn (reversed). Reliability testing of Mini Markers shows that it also has a high-reliability score compare to other inventories such as NEO-FFI. Even though Mini Markers is not as efficient to measure personalities as BFI, it has a significant validity on the determination of personalities. Considering the length of the study, Saucier's Mini Markers will be the most suitable option for the study.

Considering the period of the study, Mini Markers was decided to use for personality analysis on the thesis. Also, various research indicated that the length of psychological tests reduced the motivation of participants, and it may cause to scare participants at the beginning of the test to complete (Burisch, 1984). Thus, measuring with an abbreviated test can be considered as the best way for the thesis. In addition to that, Mini Markers have been used as a measurement tool in several investigations (e.g. Cafaro et al., 2012; Chang, 2009; Graaf and Allouch, 2014; Karampela, Tregear and Ansell, 2014; Rutjes, 2013)

In the study, the first set of questions participants faced "Rate the following adjectives considering your personality. (You can compare yourself with your peers around you.)" in order to determine participants' extroversion level. They have scored their personality through the extroversion factor adjectives of Mini-Marker which are talkative (konuşkan), extroverted (dışa dönük), bold (cesaretli), energetic (enerjik), shy (utangaç), quiet (sakin), bashful (çekingen), and withdrawn (içe dönük). The adjectives were translated according to Somer and Goldberg's (1999) study about Turkish trait-descriptive adjectives. Also, the order of all adjectives is shown randomly for each participant and a 9-point Likert scale was used to score; 1 stated as “absolutely disagree”, 9 stated as “absolutely agree”.

After watching random one of the video, to understand whether chatbots personalities were perceived as designed, extroversion scale of mini-marker adjectives was asked randomly to measure chatbot personality this time. The set of questions for the chatbot was “Rate the following adjectives considering the personality of the chatbot you have watched in the video.”. A 9-point Likert scale was used to score as before. After the rating, participants were asked to answer (the non-mandatory) open-ended questions such as “What do you think about the personality of the chatbot? What made you think of that?”

In the next set, questions are asked in order to comprehend *fun, liking, competence*, and *usefulness of interaction* dimensions of the chatbots. All sets of questions had a 9-point Likert scale, from “absolutely disagree” to “absolutely agree” and, as in all question sets, adjectives under each dimension were shown randomly as well. Adjectives that used for the category of *fun* dimension were determined according to Isbiter and Nass’s (2000) study; which are “fun”, “interesting”, “exciting”, and “satisfying”. These adjectives also categorized as *enjoyment of interaction* in Lee et al. ’s (2006) study and as *satisfaction* in Nass et al. ’s (1994) study. For *liking* dimension (Isbister and Nass, 2000) which can be also named as *social attraction* (Lee et al., 2006) or *attraction* (Nass et al., 1994), questions were “How much did you like the chatbot?” and “Would you like to interact with a chatbot like this?”. *Fun* and *liking* dimensions will be used for the determination of user preferences.

Moreover, *competence*, which can be also stated as the usefulness of the chatbot (Isbister and Nass, 2000), has a set of adjectives such as “helpful”, “clever”, “intelligent”. On the final set of questions, participants were asked for *usefulness of the interaction* dimension (Isbistar and Nass, 2000) through a trait set which contains adjectives such as “helpful” and “useful”, which is also mentioned as *benefit* dimension (Nass et al., 1994). *Competence* and *usefulness of interaction* dimensions will be evaluated to determine the perception of intelligence levels. As in all question sets, a 9-point Likert scale was used. After every dimension rating, participants were asked to answer (the non-mandatory) open-ended questions such as “What made you think of that?”

4.2 Design

There are two reasons why extroversion personality factor is the main focus of the thesis and the study. First of all, as Wilt and Revelle (2008) mentioned, extroversion can be seen as the fundamental personality type that investigated by psychology researchers (Costa & McCrae, 1992a; Digman, 1990; Eysenck and Himmelweit, 1946; Goldberg, 1990; Norman, 1963) because extroversion has a vital role to help in order to explain various behaviors (Funder, 2001). Therefore, it can be accepted as a core factor of the personality and the reason why extroversion is the factor that is investigated.

Moreover, not only in the Big Five but almost all of the personality theories have extroversion personality type in their way. For example, Eysenck Personality Theory divides personality two dimensions that are extroversion/introversion and high-neuroticism/low-neuroticism. Even though various theories do not mention literally about extroversion itself, extroversion personality types can be read. For instance, Kiesler's Interpersonal Circle has 16 personality types, which reference points of the personalities come from dominance and friendliness; and intersection range of dominant and friendly personalities contain extroverted personalities (Dryer, 1999). Another example can be Cattell's sixteen factors of personality that contains traits of extroverted personality such as outgoing, dominant, self-sufficient. Likewise, introvert personality traits like reserved, submissive, group dependent can be seen on Cattell's sixteen factors. Therefore, extroversion personality has been had immense importance on personality theories.

The second reason why the study is processed through extroversion personalities is that, in the case of conversational user interfaces, reflecting characteristics traits of extroversion or introversion is easier than the other personality factors. It is not so specific and contains vary of traits can be examined in the design of the chatbot. It does also not so widen and can be distinguished from another type of factors such as agreeableness, conscientiousness, neuroticism, and openness to experience. Moreover, the most observable and correctly transmitted (Lippa and Dietz, 2000) and quick recognizable (Kammrath, Ames and Scholer, 2007) factor of Big Five is considered as extroversion.

According to Goldberg's the Big Five Factor markers (1992) extroverted personalities described with traits as extroverted, talkative, assertive, verbal, energetic, bold, active, daring, vigorous, and unrestrained whereas introverted ones with introverted, shy, quiet, reserved, untalkative, inhibited, withdrawn, timid, bashful and unadventurous.

The question is how these traits can be used for designing a chatbot personality? What are the key points? In the thesis, not only words that are chosen but also other interaction parameters will be considered for the personalities. First of all, it cannot be ignored the power of the tone of voice in the conversational design. On the other hand, the visual design of the chatbot avatar, usage of emojis, usage of expressions, option suggestions, frequency of answers, willing to having a conversation, and leading the user will be used as critical factors for designing extroverted and introverted chatbots.

4.2.1 Tone of voice

Tone of voice is considered as a critical issue in conversational design. Interaction between user and assistant agent is established through words or voice. Whatever an agent uses as a communication tool, the way of talking only matters. Cummings (2017) describes tone of voice as, how you express yourself, not as what you express. It can be interpreted as the way of representing chatbot personality itself through words.

In the Nielsen and Norman Group website, it is declared that four elements are considered for tone of voice (Meyer, 2016). It is stated as "4 dimensions of tone of voice", which are listed as "humor", "formality", "respectfulness", and "enthusiasm". Moreover, the bipolar scale of these dimensions is shown as "funny/serious", "formal/casual", "respectful/irreverent", and "enthusiastic/matter-of-fact", in order (Meyer, 2016). When all dimensions are addressed to the markers of the Big Five (Goldberg, 1992), it can be seen that humor and respectfulness dimensions of the tone of voice match to the agreeableness factor; while formality can be addressed to conscientiousness factor according to the markers related with it. Moreover, the enthusiasm dimension of the tone of voice match to the extroversion factor.

Therefore, in order to measure the extroversion difference of chatbots, three of the four tone of voice dimensions should be the same level in two chatbots and the enthusiasm dimension should be arranged and applied on the chatbots, which means that the extroverted chatbot should have a more enthusiastic tone of voice while the introverted chatbot is being more matter-of-fact.

Besides Meyer’s (2016) declaration, in the case of extroverted and introverted chatbots, the extroverted chatbot that is used in the study should have a sharper word chose on the answers while the introverted one has a more modest way of talking. As Nass et al. (1995) applied in their investigation, dominant personalities have a high confidence level than submissive ones. Since dominance is also one of the markers of extroverted personalities (Goldberg, 1992), in this study, the extroverted chatbot should answer questions with self-confidence as Nass et al. (1995) did on their research.

For this aim, the introverted chatbot consults to the user before the suggestion. On the contrary, the other one makes decisions by itself and invite the user. For instance, the extrovert says “Compared to the previous month, I see an increase in your clothing exchange. I have a suggestion for you!”. In the introverted one, this issue happens by asking, “Can I suggest to you about your spendings?”.

Moreover, the extroverted one uses more “we” or “I” as a subject where the introverted one uses more “you” in as a sign of being a personality that does not take responsibility (Gill and Oberlander, 2002). For example, the extroverted chatbot says that “Which card of yours shall we look at for the spendings?”; whereas, the introverted one mention it by saying, “Which card of yours would you like to look at for your spendings?”.

4.2.2 Chatbot Avatar

In the profile picture of the chatbots, it was decided to use a robot figure as an avatar of the chatbots because using the women or men figure may direct participants observations and make their decisions under the bias that are socially constructed; such as that masculine figures have more extroverted personalities than feminine ones (Bem, 1974). There are; however, some different points on the pictures depend on their personalities.

The main figure of the avatar was kept still, but some elements that build the picture was designed according to the personalities. It is stated that extroverted personalities prefer colors with more saturated hues and high color contrast, and bold lines, and sharp edges on the shapes rather than the introverted personalities. Contrast to the extroverted personality, people who have introverted personalities prefer desaturated colors, green hues, and thin or rounded shapes (Karsvall, 2002). Choungorian's research (1967) shows that introverted people like more cold color tones such as green or blue while extroverted people prefer warm colors like red or yellow.

Avatars of the study are designed according to the investigation results about visual design preferences of the personalities. The introverted chatbot has a desaturated blue color. This chatbot has rounded corners and thin lines on eyes and mouth. While determining the color of the extroverted chatbot, the blue hue in the color of the introverted chatbot was decreased to the lowest level, and the red hue in color was increased. Thus, the orange color was obtained. Also, the saturation level of the extroverted chatbot was set to +30, while the level of the introverted chatbot was maintained at -30. In addition to that, the extroverted chatbot avatar has more bold lines, sharp edges than the other avatar (Figure 4.1).

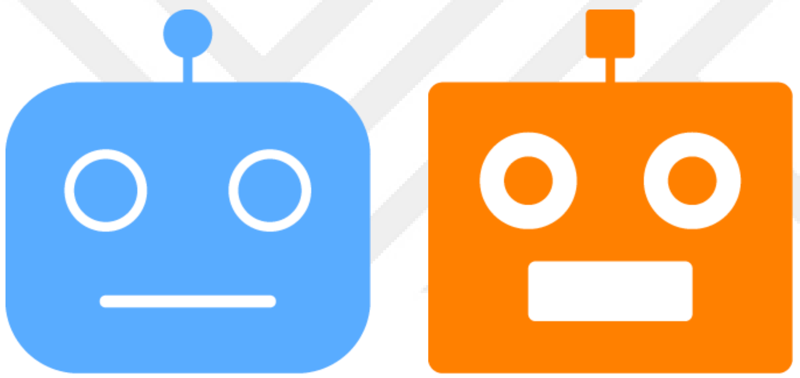

Figure 4.1 : The Introverted (on the left) and Extroverted Chatbot (on the right) Avatars.

4.2.3 Expressions

Expressing the emotions easily indicates to have an extroverted personality (Freyd, 1924). Therefore, the extroverted chatbot uses expressions on its conversation such as “Of course!”, “Hmm”, and “Super!”. These expression words were used to support the personality. Contrast to the extroverted chatbot, the introverted chatbot does not use extra words in the conversation.

4.2.4 Emojis

Emojis can also be evaluated as a specialized version of expressions for online conversations. Therefore, it can be said that emojis are also used as an easy way of conveying feelings like expressions. Moreover, extroverted people are also described as personalities that express their feelings more comfortable than introverted personalities. Therefore, the usage of emojis on the chatbots is inevitable in the study. The extroverted chatbot uses emojis as much as it can on the conversation while the introverted one does not use emojis because introverted people can be described as personalities that do not express their feelings to the outside.

Moreover, in the usage of emojis, it was aimed to be respectful, warm, and kind. The conversation should not seem as reckless or arrogant, because only the extroversion factor on the chatbots would like to be changed on the study and agreeableness factor, which contains characteristics such as warmness, kindness, respectfulness, is intended to be kept stable on both chatbots. For this aim, the extroverted chatbot uses emojis, which are named as “smiling face”, “robot face”, “waving hand”, “eyes”, “backhand dimension pointing down”, and “woman dancing” (Emojipedia, 2019).

4.2.5 Option buttons

Today, most of the chatbots suggest option buttons after their questions for leading users' answers. This method generally uses on the less smart chatbots because this kind of chatbots need to take the answer with exact words that they know. Otherwise, they do not comprehend what users tell them. There is no difference in the intelligence of the chatbots used in the study. However, the extroverted chatbot uses option button on the conversation in order to give the user options and lead the user on their answers. The other chatbot, moreover, does not use option buttons on the conversation; thus, the introverted chatbot does not affect the user's responses and can be seen as a quieter, introverted and timid chatbot.

4.2.6 Amount and frequency of answers

Besides the other parameters mentioned above; the amount and frequency of answers are also an effective way of chatbot personalities. Answers of chatbots are given through bubbles as a tool of an online platform. Bubbles’ amount refers to how much bots talk and express their emotions and thoughts. Therefore, extroverted one’s amount of answers are much more than the other, and as the number of bubbles increases, this situation also affects the frequency of sending.

Moreover, the frequency of sendings and speed of these answers can be referred to how energetic the bots are; so, extroverted one’s frequency is more in order to emphasize the extroversion of the bot. For example, the total sending amount of the extroverted chatbot's is 23, whereas introverted one's is 11 bubbles. As a summary, on the extroverted one; the amount and frequency of the bubbles are more than the introverted one. Since not only setting the tone of voice but also designing interactions on the conversation is an effective way to create personality in chatbots.

4.2.7 Willing to having a conversation

As a part of being a talkative, energetic, and enthusiastic chatbot, the extroverted chatbot should have to more willing to have a conversation with users compare to introverted one because introverted chatbot's personality should be more shy, quiet, reserved. That is why, it is considered that extroverted chatbots do not tend to finish a conversation, on the contrary, they try to continue by asking new questions or suggesting new options to users; whereas, introverted chatbots tend to be more duty-oriented, not asking further questions for chatting.

For example, on the extroverted chatbot start the conversation before the user; and at the end of the conversation, the chatbot tries to generate a new discourse by asking if she needs support in another matter. On the contrary, the introverted one does not start the conversation at the beginning and do not ask additional questions to the user. At the end of the transaction, the introverted chatbot, in contrast to extroverted one, do not ask for if the user needs anything else.

It will be easier to comprehend the chatbots' conversations through scenarios. Therefore, the next section will examine how the different types of chatbots follow different paths in the same scenario.

4.3 The Scenario of The Videos

The research aims to evaluate how the users evaluate the chatbots in different personalities and whether the user personalities are responsible in these evaluations. For this purpose, there are two different scenario-based videos on the study in order to show participants. Every participant watches one of the videos, the scenario with the extroverted chatbot or the introverted one. In other words, the between-subject method is used in the research.

The scenario of the videos is about the same topic that can be named as financial support. Participants watch a conversation about finance between user and chatbot. The flow contains four major conversations such as welcoming, getting information about account balance, making a transaction for spendings, and greetings. All flow and information provided are the same in both chatbots. How this information is delivered depends on the personality of the chatbot. For example, the extroverted chatbot uses so many emojis than the introverted one which does not use emojis in order to show that the extroverted one is more energetic and more enthusiastic than the introverted one. Also, the extroverted one uses expressions to be more extrovert its emotions like "hmm", or "super!".

Moreover, the extroverted one suggests options and expresses its thoughts, whereas the introverted chatbot does not give any options to the user or say anything extra. The reason behind that is extroverted personalities more talkative and more willing to have a conversation than the introverted ones. Separately explaining all parts of the flow can be more illustrative for the videos; therefore, four steps of the conversations will be examined next.

First, welcoming differs in the chatbot, for example; the extroverted chatbot is a proactive bot by welcoming the user before users and introduce itself, and finally ask for the purpose of the user for visiting. The introverted one waits for the user to begin the conversation. After the user's "hello", the chatbot also welcomes the user and ask directly for what kind of service the user would like to have, without introducing itself.

The second part of the conversation is to getting information from the chatbots. The user wants the learn account balance of the last month in a specific card from the chatbots. In both video, the user asks for "I want to learn my spendings". There are two information needed to transact the information of spendings; first is spendings of which card, second is spendings between which dates. Even if the introverted one should be more reserved and make no suggestion to the user; for the first information to get, both chatbots use card buttons. This is the only place in the case that the introverted chatbot uses an option button since there is no other possibility to get this information from the user. However, the rest of the introverted chatbot questions does not suggest an option to user in order to be more shy, bashful, and withdrawn. For the date information of the spendings, the introverted one does not offer options for the question and wait for the user to think the date range. The extroverted bot suggests date ranges to the user like the rest of the conversations. In fact, in the case of the date range for spendings, the extroverted one declare its idea for the most suitable option by saying that “what would you say to look at the last month spendings?”.

The third part of the flow continues with the proposal for a transaction by the chatbots. The proposed issue is to notify the user by sending a notification if the limit is exceeded in order to make their payments more under controlled. The challenge for the introverted chatbot is that the scenario itself has an extroverted attitude. When thinking about the duty of a chatbot, which is providing a service for customers’ sake, it is an inevitable situation to make an offer even for an introverted chatbot. Therefore, the introverted one makes the proposal by its way; which means that the bot makes the suggestion by being shy, timid, and bashful. The introverted chatbot first asks the user whether it can make a suggestion for the expenses or not. The extroverted chatbot, besides, makes the offer without any permission and declare its idea about expenses and makes a suggestion.

Finally, the greeting part takes place in the flow. The extroverted one asks the user whether she needs any support for another issue after the transaction. After thanking, the chatbot also thanks to the user and notify the user that it is there for support at any time. The introverted bot does not say anything extra after the process. After thanking, the chatbot also thanks to the user and the conversation ends (APPENDIX A and B).

4.4 Participants

The study was conducted online and shared in Garanti BBVA Bank and Istanbul Technical University undergraduate, graduate mail groups, and social media such as Facebook and Twitter. The study was also conducted in Turkish with native Turkish speakers to avoid linguistic and cultural variables.

Before the videos participant were completed Mini Marker Personality Inventory for Extroversion factor in order to determine their extroversion rate. Total of 207 volunteer participants completed the questionnaire. In the personality questionnaire, participants who have a mean score below five coded as an introvert, above five coded as an extrovert. 165 participants were categorized as extrovert, 42 participants were introvert according to their responses to the inventory. After the determination of personality types, participants randomly matched to one of the videos and answered the questions about the chatbot. Since it was not technically possible to match videos according to the participant personality, an unbalanced distribution was observed in the sub-groups of chatbot and participant personalities (EC-EP, EC-IP, IC-EP, IC-IP) and 20 individuals from each subgroup were randomly selected while analyzing the study to equalize this distribution. In other words, 40 introverted participants and 40 extroverted participants were included in the analysis. In the analysis and result section, features and selection of participants will be explained more detailed.

4.5 Procedure

The test was conducted in 2x2 balanced between-subject design rather than within-subject. This method is chosen in order to be an unbiased assessment. Otherwise, participants can make an inference about the personalities of the chatbots by comparing the two videos. Extroverted and introverted participants were randomly matched to one of the videos. The participants were asked for questions about the chatbot personality and experience in the video they watched in the next stage.

5. ANALYSIS AND RESULT

In the findings and evaluation of the research, descriptive statistics about the participants, standard deviation and mean values for the personalities of the participants and the chatbot they watched, the results of the reliability analysis of all the questions and each dimension that are formed according to the literature, and the results of the ANOVA tests are explained.

5.1 Descriptive Statics About Participants

In this thesis, descriptive statistic results about the participants' characteristics are shown below (Table 5.1).

Table 5.1 : Descriptive Statistics

Descriptive Statistics N=80	**Frequency**	**Valid Percent**
Gender		
Female	48	60
Male	32	40
Age		
18-25	13	16.3
26-33	54	67.5
34-41	9	11.3
42 +	4	5
Frequency of Communicating with a Chatbot		
None	21	26.3
Once	8	10
Several Times	27	33.8
Sometimes	22	27.5
Often	2	2.5

While 60% of the 80 participants (48 people) were women, 40% (32 people) were men. 16.3% of the participants (13 people) were in the age range of 18-25, 67.5% (54 people) in the 26-33 age range, 11.3% (9 people) in the 34-41 age range, 5% (4 persons) in the age range of 42 and older. In this case, it can be said that the majority of the participants were young.

When the frequency of communicating with a chatbot was examined, it was observed that 33.8% (27 people) of participants used several times, 27.5% (22 people) used it sometimes, 26.3% (21 people) used none, 10% (8 people) once used, 2.5% (2 people) used often.

5.2 Results of Reliability Analysis

The purpose of reliability analysis is to examine the internal consistency of the concepts studied in the study. In other words, it shows the extent to which the questions in the survey reflect the research topic. While 0.60 shows that the survey is reliable, a value above 0.80 indicates that the questions in the survey are highly reliable (Kalaycı, 2014, s.404-405). The reliability analysis was applied to all expressions/questions which were measured with a 9-point Likert scale. The calculated overall reliability coefficient is Cronbach's Alpha = 0.893. Since this value is greater than 0.80, it can be said that there is a very reliable questionnaire (Table 5.2).

Table 5.2 : Reliability Analysis Results of All Statements

Cronbach’s Alpha Coefficient	Variable Number (N)
0.893	27

5.3 Arithmetic Mean and Standard Deviation Values of Research Variables

In the study, a 9-point Likert questionnaire was designed to determine how chatbots are perceived by participants. 1 = strongly disagree, 9 = strongly agree that participants were asked to rate statements. In the questionnaire, there are a total of 27 statements (questions), that eight statements to determine the personality traits of participants, eight statements to determine the personality traits of the chatbot that participant watched, and 11 statements to determine the traits of the chatbot. The means and standard deviation values for the statements are shown (Table 5.3).

Table 5.3 : Mean and standard deviation values of research variables.

Statements		Mean	Std. Dev.
Participant Personality Traits			
q1	Bold (Cesaretli)	5.88	1.905
q2	Withdrawm (İçe dönük)	5.06	2.415
q3	Talkative (Konuşkan)	5.38	2.263
q4	Bashful (Çekingen)	4.53	2.365
q5	Energetic (Enerjik)	5.84	1.879
q6	Shy (Utangaç)	4.38	2.172
q7	Extroverted (Dışa dönük)	5.30	2.236
q8	Quiet (Durgun)	4.49	2.210
Chatbot Personality Traits			
q9	Bold (Cesaretli)	6.63	1.709
q10	Withdrawn (İçe dönük)	2.55	1.590
q11	Talkative (Konuşkan)	7.04	1.595
q12	Bashful (Çekingen)	2.56	1.882
q13	Energetic (Enerjik)	6.48	1.876
q14	Shy (Utangaç)	2.63	1.767
q15	Extroverted (Dışa dönük)	6.98	1.518
q16	Quiet (Durgun)	2.64	1.904
Dimension Adjectives			
q17	Fun (Eğlenceli)	4.68	2.226
q18	Clever (Akıllı)	6.69	1.940
q19	Satisfying (Yeterli)	6.70	1.958
q20	Exciting (Heyecan verici)	5.01	2.184
q21	Intelligent (Bilgili)	6.66	1.862
q22	Interesting (Enteresan)	4.86	2.127
q23	Helpful (Yardımsever)	7.20	1.810
q24	Helpful (Yardımcı)	7.23	1.743
q25	Useful (Kullanışlı)	6.56	2.116
q26	How much did you like the chatbot?	6.39	1.634
q27	Would you like to interact with a chatbot like this?	5.84	2.196

In Table 5.3, it is seen that the feature with the highest average among the participants' personality traits is "bold" (Mean = 5.88), while the lowest average is "shy" (Mean = 4.38). When the statements about the personality traits of the participants are examined, it can be said that the distribution of extroverted characteristics (bold, talkative, energetic, extroverted) is positive but close to neutral. Likewise, it can be said that the distributions of introverted (withdrawn, bashful, shy, quiet) are negative but close to neutral. In this case, it can be said that the participants define themselves as neither too extroverted nor very introverted.

In this study, extroverted and introverted chatbots are designed. The chatbots were randomly monitored and asked the participants to evaluate in the questionnaire. The questionnaire was reached to the participants online using convenience sampling method. A total of 207 people participated. Only one of the chatbots were randomly displayed to participants without knowing the personalities of the participants. As a result of the field study, the number of introverted people following the introverted chatbot was 22, while the number of extroverted people was 67. The number of introverted people following the extroverted chatbot is 20, while the number of extroverted people is 98. It was found appropriate to select 20 people from each group because an equal number of participants from each group would provide an equal distribution in the sample. The selection process was obtained by a random selection of 20 people from each category. The personality characteristics of the participants and the perceived chatbot personality characteristics were obtained according to the mean of the participants' answers to eight questions. Later, these variables are converted to non-metric form since they will be used in hypothesis testing. For the participants' personality traits variable, the ones below the mean of five are coded as an introvert, the ones above five were coded as an extrovert, and the ones five are coded as neutral. For the perceived chatbot personality variable, the ones below the mean of three are coded as more introverted, the ones between three and five are introvert, the ones between five and seven are extrovert, and the ones seven and above seven are encoded as more extrovert. Accordingly, 6.3% of participants (5 people) perceived the chatbot that is monitored as an introvert, 43.7% (35 people) as an extrovert, and 50% (40 People) as more extrovert. None of the participants perceive any chatbot as more introverted.

5.4 Establishing Dimensions of Research Variables

In the study, the traits of chatbots were examined under the dimensions of *fun, competence, usefulness of interaction,* and *liking* depending on the related literature. The dimension of *competence* from the study of Nass (1995); *fun, usefulness of interaction*, and *liking* dimensions were included from Isbister and Nass's (2000) study. Since the equivalent of many traits used in English coincides with a single word in Turkish, some of the traits used in dimension measurement are reduced by taking expert opinion. These dimensions (*fun, competence, usefulness of interaction, liking*) are also based on the mean values of the traits under each dimension. The results of mean, standard deviation and reliability analysis of the information and dimensions related to the traits of the dimensions are given (Table 5.4).

Table 5.4 : Standard deviation and reliability analysis results of the trait dimensions.

Dimension	Items	Mean	Std. Deviation	Reliability (Cronbach's Alpha)
Fun	Fun (Eğlenceli) Satisfying (Yeterli) Exciting (Heyecan verici) Interesting (Enteresan)	5.31	1.619	0.758
Competence	Clever (Akıllı) Intelligent (Bilgili) Helpful (Yardımsever)	6.85	1.567	0.787
Usefulness of Interaction	Helpful (Yardımcı) Useful (Kullanışlı)	6.89	1.785	0.821
Liking	How much did you like the chatbot? Would you like to interact with a chatbot like this?	6.11	1.782	0.821

According to the reliability analysis of the dimensions, it is seen that the reliability level of each dimension is high (over 0.70). In this context, it can be said that the characteristics of each dimension are sufficient to explain the related dimensions.

Generally, the tendency towards the perception of the chatbots is positive. While participants emphasized the most *usefulness of interaction* dimension, they paid attention the least to *fun* dimension of the chatbots.

5.5 Hypothesis Tests on the Personalities and Trait Dimensions of Chatbot

Variance analysis (ANOVA) is used to test the hypotheses about whether there are differences between two or more group means. Variance analyzes vary according to the dependent variable and the number of independent variables (Kalaycı, 2014, p. 131). When there is a dependent and an independent variable, it is called One Way ANOVA, when it is a dependent and two or more independent variables, it is called Factorial ANOVA (İslamoğlu and Alnııkık, 2014, p. 309). One way Anova is a test that determines whether there is a difference between the averages in the dependent variable relative to the subgroups in the independent variable. The factorial ANOVA is a test that determines whether the two independent variables have distinct effects on the dependent variable averages as well as their common effects. While the dependent variables are the variables defined in the metric feature, the independent variables must be measured in non-metric (nominal or ordinal) features. In ANOVA, it is necessary to look at the equality of group variances. If the group variances are not equal, Welch and Brown-Forsythe test are used instead of the F test (Kalaycı, 2014, p. 132,136).

5.6 Research Hypothesis

5.6.1 Hypothesis 1: Participants will be able to recognize chatbots' personalities

One Way ANOVA test was used to test this hypothesis. While the dependent variable is the personality characteristics of the chatbot perceived by the participants, the independent variable is the personality of the designed chatbot. The results of the analysis are shown below, in Table 5.5 and Table 5.6.

When Table 5.5 is examined, it is seen that 40 extroverted chatbots (ECs) and 40 introverted chatbots (ICs) have been defined. When the mean of the chatbot personality traits was evaluated, the participants perceived the IC as an extroverted (mean = 6.66), the EC as a more extroverted (mean = 7.53) (Table 5.5). Based on descriptive statistics, a general interpretation can be made; however, in order to understand whether this situation is statistically significant, it is necessary to look at the significance levels (sig. value).

Table 5.5 : Mean and standard deviation values of the perceived chatbot personality according to the designed chatbots.

Chatbots	N	Mean	Std. Deviation
Introvert	40	6.66	1.250
Extrovert	40	7.53	1.267
Total	80	7.09	1.326

Since the p-value of the group variances that is one of the assumptions of variance analysis (ANOVA) is greater than 0.05 (p = 0.984), the group variances are distributed homogeneously. For this reason, the F test was used in the analysis of variance.

In Table 5.6, it is seen that chatbot personalities are effective in evaluating the perceived chatbot personalities (sig. = 0.003 <0.05) (Table 5.6). In other words, the EC is perceived as extroverted, while the IC is perceived as more introverted than the extroverted chatbot. The first hypothesis is accepted.

Table 5.6 : ANOVA table of perceived chatbot personality according to designed chatbot personality.

Percieved Chatbot Personality

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	15.199	1	15.199	9.591	.003
Within Groups	123.607	78	1.585		
Total	138.806	79			

5.6.2 Hypothesis 2: Personalities of chatbots will not affect perception of chatbots' capability.

One Way ANOVA test was used to test these hypotheses. While the dependent variable is the chatbot trait dimension (*fun, competence, usefulness of interaction, liking*), the independent variable is the chatbot personalities. The findings of the analysis results are shown in Table 5.7 and Table 5.8 below.

In Table 5.7, it is seen that participants have defined 40 extroverted chatbots and 40 introverted chatbots. About the participants' perceptions on personality trait dimensions, the IC's *fun* level was found neutral (mean = 5.05); the EC's fun level was evaluated positively close to the neutral (mean = 5.58). It means that the IC is perceived as less fun than the extroverted one. In *competence* dimension, the IC's competence was evaluated positively (mean. = 6.81) while the EC's competence was also considered the chatbot to be positively (mean = 6.89). In *usefulness of interaction* dimension, the IC was evaluated as highly positive (mean = 6.78), and interaction of the EC (mean = 7.01) was considered as to be more useful than the introverted one. In the examination of *liking* dimension, the IC was considered as positive (mean = 6.28) for liking, and the IC was liked more than the EC (mean = 6.09) (Table 5.7). As a result, the EC has a higher perception in the dimensions of *fun*, *competence*, and *usefulness of interaction* than the IC while the IC has a more positive perception on *liking* dimension than the EC.

Table 5.7 : Mean and standart deviation values of the personality dimensions.

	Chatbot	N	Mean	Std. Deviation
Fun	Introvert	40	5.05	1.547
	Extrovert	40	5.58	1.666
	Total	80	5.31	1.619
Competence	Introvert	40	6.81	1.516
	Extrovert	40	6.89	1.635
	Total	80	6.85	1.567
Usefulness of interaction	Introvert	40	6.78	1.732
	Extrovert	40	7.01	1.852
	Total	80	6.89	1.785
Liking	Introvert	40	6.28	1.597
	Extrovert	40	5.95	1.957
	Total	80	6.11	1.782

A descriptive statistic can only be interpreted in a general level; however, it is crucial to examine the level of significance (sig.) to see if these differences are statistically significant.

In the homogeneity test of group variances, which are the assumptions of variance analysis (ANOVA); significance levels were found to be 0.627 in *fun*, 0.364 in *competence*, 0.604 in *usefulness of interaction*, and 0.073 in *liking* dimension. Since the significance levels (also called sig or p) are greater than 0.05, the group variances of the sizes are distributed homogeneously. Thus, it was found appropriate to perform F test for all dimensions. The results of the analysis are given in Table 5.8.

In Table 5.8, it can be said that the chatbot personalities was not significant on dimensions of *fun* (sig. = 0.078), *competence* (sig. = 0.80), *usefulness of interaction* (sig. = 0.926) and *liking* (sig = 0.75) due to the fact that the significance level of these dimensions is higher than 0.05 (Table 5.8).

Table 5.8 : ANOVA table of the personality dimensions between chatbot personality.

		Sum of Squares	df	Mean Square	F	Sig.
Fun	Between Groups	5.513	1	5.513	2.133	.148
	Within Groups	201.550	78	2.584		
	Total	207.063	79			
Competence	Between Groups	.139	1	.139	.056	.813
	Within Groups	193.883	78	2.486		
	Total	194.023	79			
Usefulness of Interaction	Between Groups	1.128	1	1.128	.351	.555
	Within Groups	250.719	78	3.214		
	Total	251.847	79			
Liking	Between Groups	2.113	1	2.113	.662	.418
	Within Groups	248.875	78	3.191		
	Total	250.988	79			

In other words, the chatbot personalities observed by the participants do not have a distinctive effect on the evaluation of these personality trait dimensions, and all dimensions are evaluated at the same level. *Competence* and *usefulness of interaction* are trait dimensions created to determine the perceived intellectual levels of chatbots, and Hypothesis 2 has been accepted since there is also no difference in these dimensions, such as other dimensions.

5.6.3 Hypothesis 3: The personalities of the participants will affect their perception.

If the participants' personalities have an impact on perceived chatbot personalities, the factorial ANOVA should be used for the variance analysis models (ANOVA). The results of the analysis, in which the personalities of the participants and chatbots are determined as independent variables and the perceived chatbot personalities as dependent variables are shown in Table 5.9 below.

As shown in Table 5.9, the introverted participants (mean = 6.73) who watch the IC perceived the chatbot more introverted than the extroverted participants (mean = 6.58). The extroverted participants (mean = 7.76), who faced the EC, perceived the chatbot more extroverted than introverted participants (mean = 7.30) (Table 5.9).

Table 5.9 : Mean and standart deviation values of perceived chatbot personality according to chatbot and participant personalities.

Dependent Variable: Perceived Chatbot Personality

Chatbot Personality	Participant	Mean	Std. Dev.	N
Introvert	Introvert	6.73	1.463	20
	Extrovert	6.58	1.028	20
	Total	6.66	1.250	40
Extrovert	Introvert	7.30	1.202	20
	Extrovert	7.76	1.320	20
	Total	7.53	1.267	40
Total	Introvert	7.02	1.353	40
	Extrovert	7.17	1.310	40
	Total	7.09	1.326	80

The averages are very close to each other. The significance values (sig. values) should be evaluated to determine if these mean differences are statistically significant.

In Table 5.10, it is seen that the designed chatbot personality is significant on the perceived chatbot personality (sig. = 0.003 <0.05). It can be said that the chatbot personalities have an effect of 11.1% on the perceived personalities. Besides, it was determined that the personalities of the participants (sig. = 0.588> 0.05) and the interaction effect of the chatbot personality with the participant personality (sig. = 0.287> 0.05) were not significant on the perceived chatbot personalities (Table 5.10). In other words, in the determination of the perceived chatbot personality, the designed chatbot personality provides a distinctive effect, while the personalities of the participants and the interaction of designed chatbots with the participants do not provide a distinctive difference. As shown in Figure 5.1, it can be said that this is since the means of subgroups are close to each other (Figure 5.1). In other words, it is possible to say that the personalities of the chatbots are perceived correctly not because of the personality of the participants, but because of the personality of the chatbots.

Table 5.10 : Tests of between-subject effects.

Dependent Variable: Perceived Chatbot Personality

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squar d
Corrected Model	17.505[a]	3	5.835	3.656	.016	.126
Intercept	4025.845	1	4025.845	2522.347	.000	.971
The Chatbot	15.199	1	15.199	9.523	.003	.111
Participant	.473	1	.473	.296	.588	.004
The Chatbot * Participant	1.833	1	1.833	1.149	.287	.015
Error	121.301	76	1.596			
Total	4164.651	80				
Corrected Total	138.806	79				

a. R Squared = .126 (Adjusted R Squared = .092)

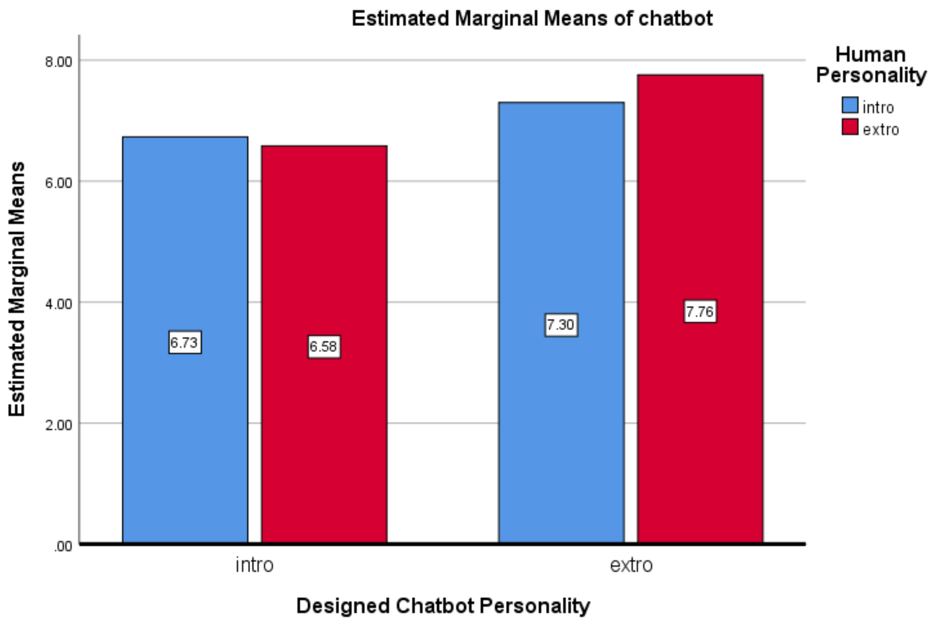

Figure 5.1 : Sub-group means of chatbot and participant personalities according to the perceived chatbot personality.

In addition to the chatbot personalities, to measure that the personalities of the participants affect the perception of trait dimensions, it is necessary to use Factorial ANOVA from variance analysis models (ANOVA). In the analysis, the personality of the participants and chatbots were determined as independent variables, whereas trait dimensions are considered as dependent variables. The results for the analysis are shown in Table 5.11 and Table 5.12 for *fun* dimension, Table 5.13 and Table 5.14 for *competence* dimension, Table 5.15 and Table 5.16 for *usefulness of interaction* dimension and Table 5.17 and Table 5.18 for *liking* dimension.

When Table 5.11 is examined, not only introverted participants (mean = 5.03) but also extroverted participants (mean = 5.08) perceived *fun* dimension of introverted chatbot neutrally. The EC's *fun* dimension was perceived as positively neutral by introverted participants (mean = 5.35) whereas it was perceived as more positively by the extroverted participants (mean = 5.80) than the introverted ones (Table 5.11). The means are very close to each other. The significance values (sig. values) should be evaluated to determine if these mean differences are statistically significant.

As shown in Table 5.12, it can be said that there is no difference between the group averages of chatbot personality in the perception of *fun* dimension (sig. = 0.151> 0.05) (Table 5.12). In other words, the extroverted or introverted chatbots do not make any significant difference in the perception of *fun* dimension. In addition to this, it was found that both the participant's personality (sig. = 0.492> 0.05) and the interaction effect of the chatbot personality with the participant personality (sig. = 0.583> 0.05) were not significant on the perception of *fun* dimension.

Table 5.11 : Mean and standart deviation values of the fun dimension according to chatbot and participant personalities.

Dependent Variable: Fun

Chatbot Personality	Participant	Mean	Std. Deviation	N
Introvert	Introvert	5.03	1.495	20
	Extrovert	5.08	1.635	20
	Total	5.05	1.547	40
Extrovert	Introvert	5.35	1.536	20
	Extrovert	5.80	1.798	20
	Total	5.58	1.666	40
Total	Introvert	5.19	1.505	40
	Extrovert	5.44	1.736	40
	Total	5.31	1.619	80

In other words, in the perception of *fun* dimension, neither the personality of the designed chatbots, the personality of the participants, nor the interaction of the personality of the designed chatbots and the participants provides a distinctive difference. As shown in Figure 5.2, this can be said to be since the subgroup means (IC&IP, IC&EP, EC&IP, EC&EP) are close to each other.

As shown in Table 5.13, *competence* dimension of the IC was perceived positively by both introverted participants (mean = 6.82) and extroverted participants (mean = 6.80). Also, *competence* dimension of the extroverted chatbot was perceived positively by both introverted participants (mean = 6.85) and extroverted participants (mean = 6.93) (Table 5.13). The means are very close to each other. The significance values (significant values) should be evaluated to determine if these mean differences are statistically significant.

Table 5.12 : Tests of between-subject effects.

Dependent Variable: Fun

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	7.563[a]	3	2.521	.960	.416	.037
Intercept	2257.813	1	2257.813	860.119	.000	.919
The Chatbot	5.513	1	5.513	2.100	.151	.027
Participant	1.250	1	1.250	.476	.492	.006
The Chatbot * Participant	.800	1	.800	.305	.583	.004
Error	199.500	76	2.625			
Total	2464.875	80				
Corrected Total	207.063	79				

a. R Squared = .037 (Adjusted R Squared = -.002)

As shown in Table 5.14, it can be said that the chatbot personality did not make a difference in the perception of *competence* dimension at the significance value of 5% (sig. = 0.816). In other words, designing a chatbot with introverted or extroverted personalities does not make any significant difference in the mean perception of *competence* dimension. In addition to this, it was found that both the participant's personality (sig. = 0.926> 0.05) and the interaction effect of the participant personality with the chatbot personality were not significant on the perception of *competence* dimension (sig. = 0.889> 0.05) (Table 5.14). In other words, in determining *competence* dimension perceptions, neither the personality of the chatbot, the personality of the participants nor the interaction effect of the participant personality with the designed chatbot personality provides a distinctive difference. As shown in Figure 5.3, it can be said that the mean of sub-group (IC&IP, IC&EP, EC&IP, EC&EP) is close to each other (Figure 5.3).

As shown, the usefulness of the interaction dimension of the IC was perceived more positively by introverted participants (mean = 7.03) than extroverted participants (mean = 6.53). Also, *competence* dimension of the EC was perceived more positively by introverted participants (mean = 7.08) than extroverted participants (mean = 6.95) (Table 5.15). The averages are very close to each other. The significance values should be evaluated to determine if these mean differences are significant.

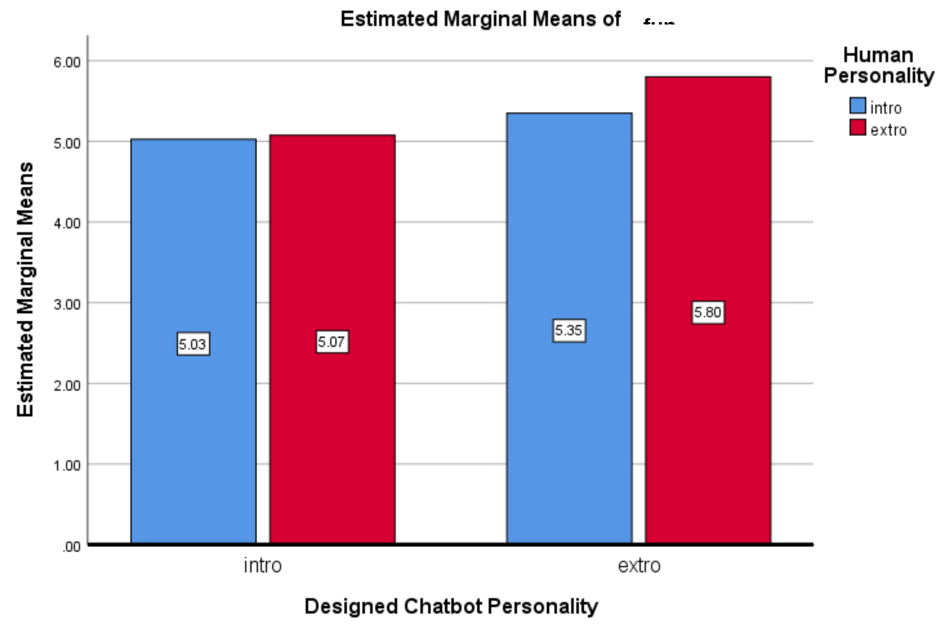

Figure 5.2 : Sub-group means of chatbot and participant personalities according to the fun dimension.

As shown in Table 5.16, it can be said that the chatbot personality did not make a significant difference in the perception of *usefulness of interaction* dimension at the significance value of 5% (sig. = 0.558) (Table 5.16). In other words, the introvert or extroverted chatbots do not make any difference in the mean of perception of usefulness of interaction. In addition to this, it was determined that participant's personality (sig. = 0.442> 0.05) and the interaction effect of the participant personality with the chatbot personality (sig. = 0.442> 0.05) were not significant on the perception of usefulness of interaction dimension. In other words, in determining the perception of usefulness of interaction dimension, neither the designed chatbot personality, the personality of the participants, nor the interaction effect of the participant personality with the designed chatbot personality provides a distinctive difference. As shown in Figure 5.4, it can be said that the means of the subgroup (IC&IP, IC&EP, EC&IP, EC&EP) are close to each other (Figure 5.4).

Table 5.13 : Mean and standart deviation values of the competence dimension according to chatbot and participant personalities.

Dependent Variable: Competence

Chatbot Personality	Participant	Mean	Std. Deviation	N
Introvert	Introvert	6.82	1.588	20
	Extrovert	6.80	1.481	20
	Total	6.81	1.516	40
Extrovert	Introvert	6.85	1.217	20
	Extrovert	6.93	2.001	20
	Total	6.89	1.635	40
Total	Introvert	6.83	1.397	40
	Extrovert	6.87	1.739	40
	Total	6.85	1.567	80

Table 5.14 : Tests of between-subject effects.

Dependent Variable: Competence

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	.212[a]	3	.071	.028	.994	.001
Intercept	3753.800	1	3753.800	1471.997	.000	.951
Designed Chatbot	.139	1	.139	.055	.816	.001
Participant	.022	1	.022	.009	.926	.000
Designed Chatbot * Participant	.050	1	.050	.020	.889	.000
Error	193.811	76	2.550			
Total	3947.823	80				
Corrected Total	194.023	79				

a. R Squared = .001 (Adjusted R Squared = -.038)

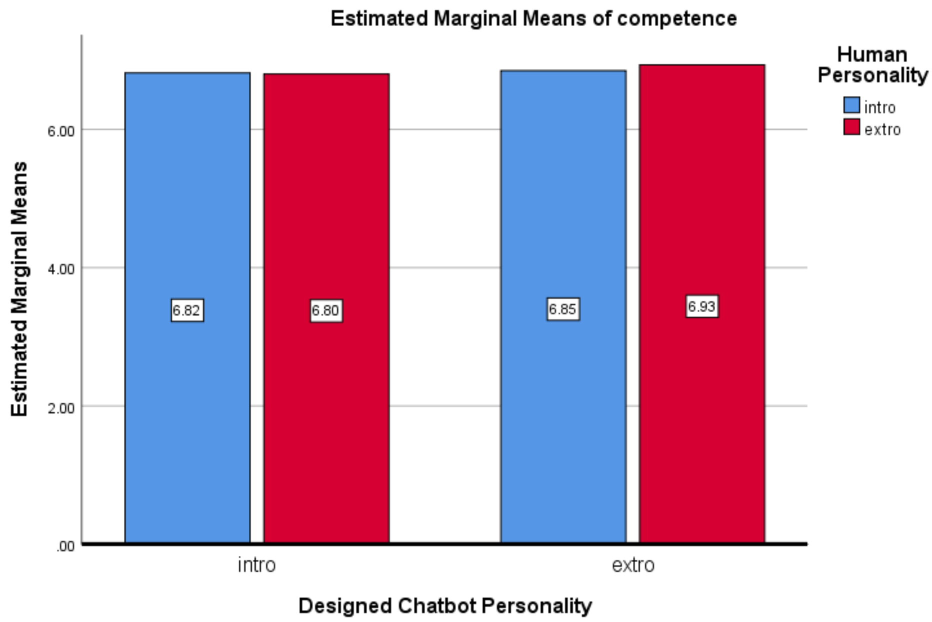

Figure 5.3 : Sub-group means of chatbot and participant personalities according to the competence dimension.

Table 5.15 : Mean and standard deviation values of usefulness of interaction dimension according to chatbot and participant personalities.

Dependent Variable: Usefulness of Interaction

Chatbot Personality	human	Mean	Std. Deviation	N
Introvert	Introvert	7.03	1.690	20
	Extrovert	6.53	1.781	20
	Total	6.78	1.732	40
Extrovert	Introvert	7.08	1.462	20
	Extrovert	6.95	2.212	20
	Total	7.01	1.852	40
Total	Introvert	7.05	1.560	40
	Extrovert	6.74	1.994	40
	Total	6.89	1.785	80

As shown in Table 5.17, not only introverted participants (mean = 6.55) but also extroverted participants (mean = 6.00) perceived *liking* dimension of the introverted chatbot positively. Also, the score of introverted participants for the dimension was higher compared to the extroverted ones. Both introverted participants (mean = 5.85) and extroverted participants (mean = 6.05) perceived liking dimension positively; however, the score of introverted participants was lower than the extroverted participants (Table 5.17). The means are very close to each other. The significance values (sig. values) should be evaluated to determine if these mean differences are statistically significant.

Table 5.16 : Tests of between-subject effects.

Dependent Variable: Usefulness of Interaction

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	3.784[a]	3	1.261	.386	.763	.015
Intercept	3801.903	1	3801.903	1164.806	.000	.939
The Chatbot	1.128	1	1.128	.346	.558	.005
Participant	1.953	1	1.953	.598	.442	.008
The Chatbot * Participant	.703	1	.703	.215	.644	.003
Error	248.063	76	3.264			
Total	4053.750	80				
Corrected Total	251.847	79				

a. R Squared = .015 (Adjusted R Squared = -.024)

As shown in Table 5.18, it can be said that the chatbot personality did not make a significant difference in the perception of *liking* dimension at the significance value of 5% (sig. = 0.421). In other words, the introvert or extroverted chatbots do not make any difference in the mean of perception of *liking* dimension. In addition to this, it was determined that participant's personality (sig. = 0.664> 0.05) and the interaction effect of the participant personality with the chatbot personality (sig. = 0.354> 0.05) were not significant on the perception of *liking* dimension (Table 5.18).

In other words, in determining the perception of *liking* dimension, neither the designed chatbot personality, the personality of the participants, nor the interaction effect of the participant personality with the designed chatbot personality provides a distinctive difference. As shown in Figure 5.5, it can be said that the means of the subgroup (IC&IP, IC&EP, EC&IP, EC&EP) are close to each other (Figure 5.5).

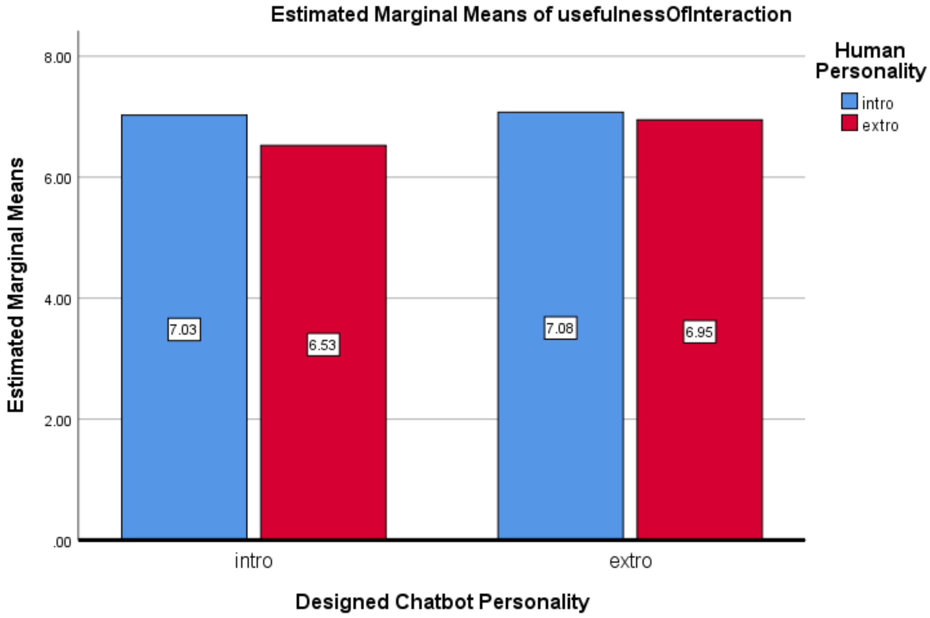

Figure 5.4 : Sub-group means of chatbot and participant personalities according to usefulness of interaction dimension.

Table 5.17 : Mean and standart deviation values of liking dimension according to chatbot and participant personalities.

Dependent Variable: Liking

Chatbot Personality	Participant	Mean	Std. Deviation	N
Introvert	Introvert	6.55	1.521	20
	Extrovert	6.00	1.662	20
	Total	6.28	1.597	40
Extrovert	Introvert	5.85	1.733	20
	Extrovert	6.05	2.200	20
	Total	5.95	1.957	40
Total	Introvert	6.20	1.648	40
	Extrovert	6.03	1.925	40
	Total	6.11	1.782	80

Table 5.18 : Tests of between-subject effects.

Dependent Variable: Liking

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	5.537[a]	3	1.846	.572	.635	.022
Intercept	2989.013	1	2989.013	925.504	.000	.924
The Chatbot	2.113	1	2.113	.654	.421	.009
Participant	.613	1	.613	.190	.664	.002
The Chatbot * Participant	2.813	1	2.813	.871	.354	.011
Error	245.450	76	3.230			
Total	3240.000	80				
Corrected Total	250.988	79				

a. R Squared = .022 (Adjusted R Squared = -.017)

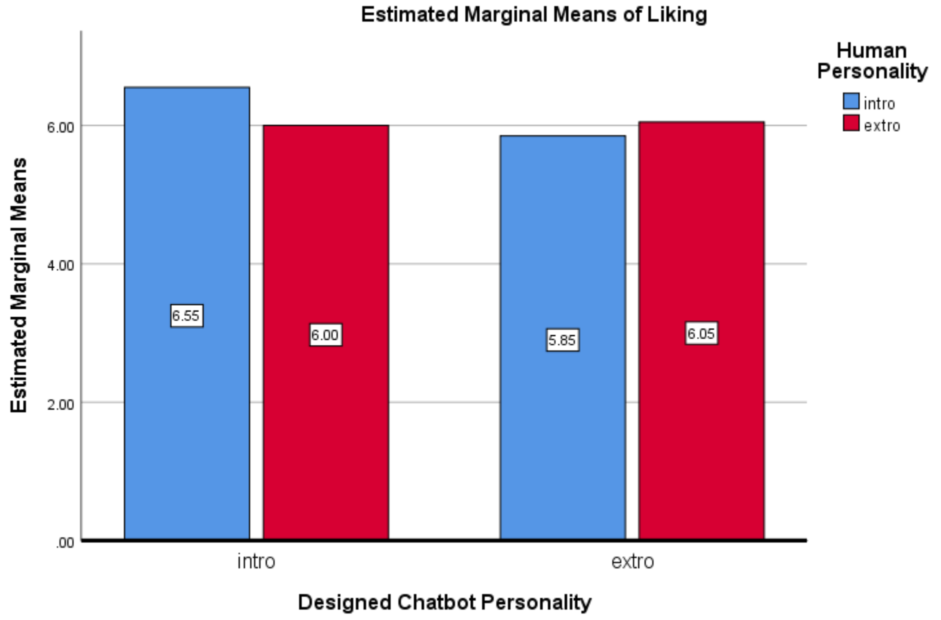

Figure 5.5 : Sub-group means of chatbot and participant personalities according to the liking dimension.

According to the results, it can be considered that Hypothesis 3, the personalities of the participants will affect their perception, has been rejected.

5.7 Open-ended Questions

In the analysis of open-ended questions for chatbot personalities, in general, the introverted chatbot was found helpful, solution-oriented, extroverted, and kind while the extroverted chatbot was found friendly, solution-oriented, energetic and talkative. Moreover, adjectives such as clever, customer-oriented, proactive, user-friendly, bashful were only used for description on the introverted chatbot; and, the adjectives such as curious, unserious, controller, energetic, and quiet were mention in order to define only the extroverted chatbot (Figure 5.6).

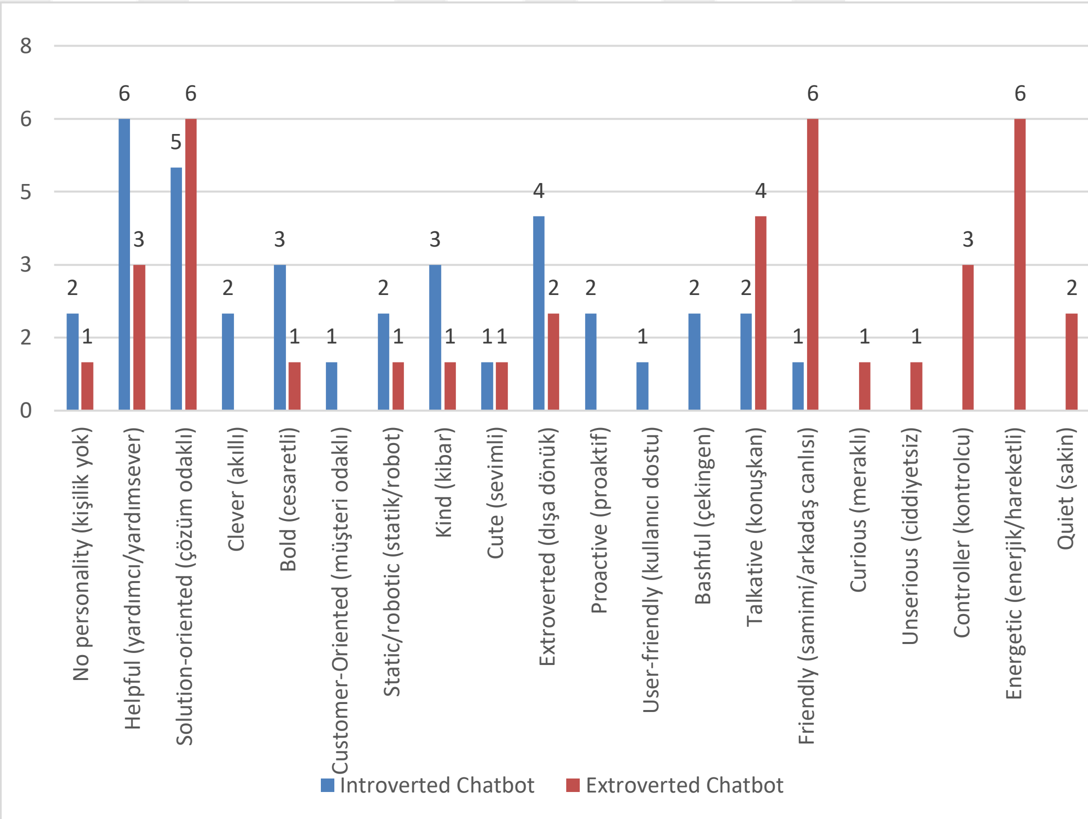

Figure 5.6 : Adjectives were used for defining personality of the introverted and extroverted chatbot.

6. DISCUSSION AND CONCLUSION

6.1 Discussion of Research Question 1: Do participants consider the extroverted chatbot as extrovert, the introverted chatbot as introvert?

According to the results of the study, it was shown that the participants perceived the IC as an extrovert, and the EC was more extroverted than the IC. Even though the IC was not perceived as an introvert, there is a significant difference between perceptions of the chatbot personalities and the IC was perceived as more introverted than the extroverted chatbot. Therefore, it can be said that Hypothesis 1 is accepted.

To understand why the IC has not been understood an introvert, the purpose of the chatbots should be examined. When the primary usage of the chatbots in daily life is overviewed, it can be mentioned that chatbots are conversation-based robots designed to make people's lives easier, allowing users to access the information they are looking for quickly and to make transactions easily. Because of the service they provide, chatbots which have introverted personality traits are not possible in a realistic scenario. The scenario in the research is designed based on a very realistic plot. First of all, both chatbots offer the user the information she is looking for and then make a suggestion. Under normal circumstances; a suggestion, which is not something that a bashful and shy person would prefer to do, was also acted by the IC in order to provide the same conditions for both chatbots. Only the IC made the suggestion more bashful, asking the user for permission before making a recommendation. Therefore, in general, the IC was perceived as extroverted, but it is more introverted compared to the EC. Moreover, the results of the research question show that design parameters that are used in the study have a significant impact on giving a personality in chatbots. In the next section, it will be discussed in details.

6.2 Discussion of Research Question 2: What are the design criteria for extroverted and introverted chatbots?

In the study, two chatbots were designed and seven design parameters were determined in order to construct a chatbot personality; which are tone of voice, chatbot avatar, expressions, emojis, option buttons, amount and frequency of answers, and willing to have a conversation. All the parameters to create a chatbot personality was suggested according to the literature; however, they needed to be tested in order to assert significant suggestions.

The results of the study show that these parameters are sufficient to create a personality in the case of extroversion. Especially use of emojis has an impact on the creation of the extroverted chatbot. In open-ended questions, when the participants were asked about the personality of the chatbot, the use of emojis was declared to be useful in interpreting the personality of the extroverted chatbot. Therefore, designers should pay attention to the use of emojis when they design a chatbot which has a personality with extroverted traits.

In the case of extroverted chatbot personality, emojis such as smiling face", "robot face", "waving hand", "eyes", "backhand dimension pointing down", and "woman dancing" were used. Different personalities may need different emojis. For instance, a chatbot with a friendly personality should use emojis like "hugging face", "heart", and "upside-down face". On the other hand, a melancholic chatbot should use emojis like "crying face", "broken heart". Therefore, every emojis should be considered in their personality context. In the context of extroversion, the emojis used in the case can be considered as a set for extroverted chatbot personality.

In this perspective, expressions which can be considered as a text-based version of emojis which are used to express thoughts and feelings of the extroverted chatbot. Like emojis, expressions can be changed according to their purpose of personality. Expressions such as "Super!" and "Hmm" used in the extroverted chatbot while the introverted chatbot did not express its feeling or thoughts. Therefore, some of the personalities may not need to use expressions, and others should use expressions in their way. For instance, melancholic characteristic chatbots can use the word "Pff" as an expression of conveying its feeling.

Moreover, in the examination of the open-ended questions, participants mostly use words such as friendly, solution-oriented, and energetic for the EC. Even though it is not at the same rate, it is seen that definitions such as friendly and solution-oriented have also been used for the IC, but the word energetic has been used only to describe the EC. The reason for this can be argued that difference in the frequency and amount of the answers. In the same scenario, the EC uses 23 chat bubbles, whereas the IC makes it do with 11 bubbles. Therefore, the EC's answer frequency is more than the IC. It can be interpreted that is the reason why the EC was described as energetic.

Moreover, personality in the physical interactions (Desmet, Nicolas and Schoormans, 2008), movements are considered as a characteristic feature of the personality. For example, the dominant characteristics of the product conducted through strong and sharp movements. In addition to that; as Lee, Peng, Jin and Yan (2006) stated, movements of the extroverted social robots are considered as more extensive and more frequent than introverted ones (Gallaher, 1992; Gifford, 1991; Isbister and Nass, 2000). Therefore, even if the medium of the interaction is changed, movement patterns can be evaluated as a characteristic feature for product personality. In the case of chatbots, it is defined as frequency and the amount of the answers can be considered as a significant parameter for extroverted personality types.

Tone of voice can also be discussed as an efficient parameter for giving personality in chatbots. The way of talking through text can be altered according to intended characteristics. In the study, the extroverted chatbot has a more confident and bold word of chose. Lee, Peng, Jin and Yan (2006) also investigated the similar tone of voice in their study for dominant interfaces. Designers should decide tone of voice of a chatbot depends on their personality intentions. Four dimensions of tone of voice, which are listed as "funny/serious", "formal/casual", "respectful/irreverent", and "enthusiastic/matter-of-fact" (Meyer, 2016) should be evaluated according to the intended personality. In extroversion, enthusiasm dimension of the tone of voice was considered; however, one or all of them can be evaluated depends on the aimed personality. For example, a chatbot with have a high agreeableness factor in its personality can be more funny than serious and more casual than formal. Therefore, designers should deliberate tone of voice of every personality in its primary purpose.

On the other hand, willing to have a conversation and option buttons the design parameters can be considered as a specific feature in the case of extrovert chatbots because it is a characteristic feature of an extrovert personality type. Other personality types also may have a variety of specified parameters in their cases. For example, a chatbot has an adventures personality may have a parameter such as willing to suggest too many new ideas. Therefore, every personality parameter should be considered in their context.

Not only the context of personality but also the tool that is used for communication can be mattered. For instance, in the case of chatbots, text-based communication method evaluated, but in voice-based user interfaces, different parameters can be involved or the same parameters can be evaluated differently in the design process. For example, in the voice user interfaces, speed or volume of the speech can be matter on the creation of the personality. Moreover, tone of voice of the assistant can be perceived differently depends on the speech tone. On the chatbots, a sentence is read by the voice of the user while the assistant has an own voice of itself in the voice-based assistants. Therefore, the media should be considered on the creation of a personality.

Overall, according to results of the study, since the EC perceived as more extroverted than the other one, it can be said that the parameters that are suggested in the thesis can be used as a tool kit for designers and designers should evaluate these parameters base on their purposed personality type and the media.

6.3 Discussion of Research Question 3: Do chatbot personalities affect the perception of chatbot capability?

The dimensions of *competence* and *usefulness of interaction* in the study were evaluated to measure whether chatbot personality affects the perception of chatbot capability. The results show that neither *competence* nor *usefulness of interaction* dimensions were not comprehended differently in the chatbots. In other words, not only the IC but also EC perceived as have a similar intelligence as intended. Therefore, it can be mentioned that chatbot capabilities do not depend on their personalities and intelligence level of the chatbots are comprehended through the capability of task accomplishments. Designers can consider giving a personality to a chatbot independently from chatbots' intelligence.

6.4 Discussion of Research Question 4: Do the personalities of the participants affect the perception of chatbot personality?

Since it is aimed to measure whether participants' personalities have an impact on their answers, the participants' personalities were also analyzed as extrovert and introvert. In the case of chatbot personality, it was shown that the perceived chatbot personalities do not vary according to the personalities of the participants. In other words, unlike other studies conducted in the field of HCl and HRI, the personality of the participants does not affect how the extroverted and introverted chatbots are perceived. Not only extroverted participants but also introverted participants perceived the EC as more extroverted while both also perceived the IC as more introverted than the extroverted one.

As a result, designers should design what kind of personality they want to reflect on a chatbot. The personalities of the target groups will not make a difference in interpreting the personalities of the chatbot. In other words, the designed chatbots will not make any difference to the user's perception depending on the personality of the users. Therefore, a chatbot should be designed according to how a personality is wanted to be presented. However, determination of the personality can be critical depends on the market the company drive. For example, a medical assistant should have a more welcoming personality and not have an immature or childish personality, while a language-teaching assistant can be more cavalier, funny, and immature. After the decision of what kind of personality should the company reflect through their chatbots, designers can apply parameters according to these personality types and should not to focus on their customers' personalities.

6.5 Discussion of Research Question 5: Do the personalities of the participants affect their perception of *fun, competence, usefulness of interaction* and *liking* dimensions of the chatbots?

In the case of participant personality impact on chatbot trait dimensions, it is shown that participant personality has no significant effect on these dimensions. In the study, four trait dimensions were measured on chatbots; *competence* and *usefulness of interaction* used to determine intelligence level; *fun* and *liking* used to understand user preferences. Without examining participant personalities, the chatbots are perceived as similarly favorable in the dimensions of *fun, competence, usefulness of interaction,* and *likin*g. It means that users think that both chatbots are fun, their intelligence level positively similar, user-chatbot interaction is satisfying, and they like both chatbots at the same level. Moreover; according to participant personalities, these thoughts are not changed. Although in the literature, there are several examples of similarity attraction between products and users; in this study, there is no significant difference between either extroverted or introverted participants' perceptions.

The reason why results of this study are different compared to the examples in the literature can be explained as the fact that this research is based on chatbot which is a newly developed technology compared to previous research. Users are expected to perform their operations successfully because many chatbots have not become as smart as presented in the study. Therefore, it is acceptable for users to be primarily functionally focused and should not pay attention to their personalities. When we look at the frequency of chatbot use of participants, it is seen that %70.1 of participants are not related to a chatbot in their daily life, who communicate with chatbots none, once, or several times in their life. Perhaps, when chatbots are evolving and becoming more involved in daily life, users will now be more careful about issues such as chatbot personality and improve their preferences in that direction.

Another reason can be that the study was conducted online; the conversations with chatbots were monitored to participants through videos, and participants did not experience the chatbots in real. If participants would interact the chatbots in their daily life routine, then, their chatbot preferences in *liking* dimension may change according to their personalities; likewise, how much users have fun with the chatbot may show changes depends on their personalities.

Consequently, chatbots are emerging technologies that might find a wider application area in the future. The personality qualities of these chatbots are a critical aspect of designing a user experience. In this study, it is indicated that extroverted and introverted personality traits can be applied to chatbots easily, and these characteristics do not impact their capability perceptions by users. In addition to that, it is shown that liking a chatbot and having fun with it does not alter to chatbot personalities, and from user to user. These outputs of the thesis may guide to designers on their journey with this new technology.

6.6 Limitations and Future Work

The limitation of the study can be seemed as conducting the study online. Although online research can help to reach more participants in a limited period, the result can be more productive with a study of interacting conversations with users and collecting data with in-depth interviews. Therefore, as future work, that kind of study can be considered.

In addition to that, different personality traits can be studied on chatbots. In this study, extroverted and introverted personality traits have been investigated; however, the same methodology can be applied with different opposite traits on chatbots. For example, dominant-submissive, formal-informal, cheerful-melancholic personality traits can be applied to chatbots. This will contribute to the development of the potential area in the literature.

Moreover, since chatbots are considered as a new developing technology, they have not found a vital place in people's lives yet. Therefore, different results can be obtained when the same study is performed at a time in the future. It will be precious to try this again in the future and to compare it with today.

7. REFERENCES

Allen, B. P., & Potkay, C. R. (1981). On the arbitrary distinction between states and traits. *Journal of Personality and Social Psychology*, *41*(5), 916.

Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological monographs*, *47*(1), i.

Al-Natour, S., Benbasat, I., & Cenfetelli, R. T. (2006). The role of design characteristics in shaping perceptions of similarity: The case of online shopping assistants. *Journal of the Association for Information Systems*, *7*(12), 34.

Aly, A., & Tapus, A. (2015, September). Multimodal adapted robot behavior synthesis within a narrative human-robot interaction. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 2986-2993). IEEE.

Apple (2019). Retrieved from https://www.apple.com/siri/

Bates, J. (1994). The role of emotion in believable agents. *Communications of the ACM*, *37*(7), 122-125.

Baumgarten, F. (1933). Die Charktereigenschaften'. [The character traits]. *Beitraege zur Charakter-und Persoenlichkeitsforschung (Whole No. 1). A. Francke, Bern, Switzerland*.

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of consulting and clinical psychology*, *42*(2), 155.

Bevacqua, E., De Sevin, E., Pelachaud, C., McRorie, M., & Sneddon, I. (2010, March). Building credible agents: Behaviour influenced by personality and emotional traits. In *Proceedings of International Conference on Kansei Engineering and Emotion Research*.

Botwin, M. D., & Buss, D. M. (1989). Structure of act-report data: Is the five-factor model of personality recaptured? *Journal of Personality and social Psychology*, *56*(6), 988.

Burisch, M. (1984). Approaches to personality inventory construction: A comparison of merits. *American Psychologist*, *39*(3), 214.

Business Insider (2017a). *Here's how chatbot metrics differ from traditional apps*, BI Intelligence, Available: www.businessinsider.com/heres-how-chatbot-metrics-differ-from-traditional-apps-2016-10, accessed 24th of June 2017.

Byrne, D., Clore, G. L., & Smeaton, G. (1986). The attraction hypothesis: Do similar attitudes affect anything?

Cafaro, A., Vilhjálmsson, H. H., Bickmore, T., Heylen, D., Jóhannsdóttir, K. R., & Valgarðsson, G. S. (2012, September). First impressions: Users' judgments of virtual agents' personality and interpersonal attitude in first encounters. In *International conference on intelligent virtual agents* (pp. 67-80). Springer, Berlin, Heidelberg.

Call-Cummings, M. (2017). Establishing communicative validity: Discovering theory through practice. *Qualitative Inquiry*, *23*(3), 192-200.

Cattell, R. B. (1943). The description of personality: Basic traits resolved into clusters. Journal of Abnormal and Social Psychology, 38, 4 7 6 - 506

Cattell, R. B., Eber, H. W., & Tatsvoka, M. M. (1970). Handbook for the sixteen-personality factor questionnaire. *Illinois: Institute for Personality ana Ability Testing.*

Chang, B. I. (2009). *Examining links between diversity and outcomes in work groups: Effects of different levels of diversity and social networks* (Doctoral dissertation, University of Akron).

Chaplin, W. F., John, O. P., & Goldberg, L. R. (1988). Conceptions of states and traits: dimensional attributes with ideals as prototypes. *Journal of personality and social psychology*, *54*(4), 541.

Chatbot 2019 Trends and Stats with Insider Reports. (2019). Retrieved from https://chatbotslife.com/chatbot-2019-trends-and-stats-with-insider-reports-fb71697deee4

Conley, J. J. (1985). Longitudinal stability of personality traits: A multitrait–multimethod–multioccasion analysis. *Journal of personality and social psychology*, *49*(5), 1266.

Costa, P. T., & McCrae, R. R. (1985). The NEO personality inventory.

Costa Jr, P. T., & McCrae, R. R. (1995). Domains and facets: Hierarchical personality assessment using the Revised NEO Personality Inventory. *Journal of personality assessment*, *64*(1), 21-50.

Choungourian, A. (1967). Introversion—extraversion and color preferences. *Journal of Projective Techniques and Personality Assessment*, *31*(4), 92-94.

Dautenhahn, K., Ogden, B., & Quick, T. (2002). From embodied to socially embedded agents–implications for interaction-aware robots. *Cognitive Systems Research*, *3*(3), 397-428.

de Raad, B., di Blas, L., & Perugini, M. (1998). Two independently constructed Italian trait taxonomies: Comparisons among Italian and between Italian and Germanic languages. *European Journal of Personality*, *12*(1), 19-41.

Desmet, P. M., Nicolas, J. C. O., & Schoormans, J. P. (2008). Product personality in physical interaction. *Design Studies*, *29*(5), 458-477.

Digital 2019: Global Digital Overview. (2019). Retrieved from https://datareportal.com/reports/digital-2019-global-digital-overview

Digman, J. M., & Takemoto-Chock, N. K. (1981). Factors in the natural language of personality: Re-analysis, comparison, and interpretation of six major studies. *Multivariate behavioral research*, *16*(2), 149-170.

Digman, J. M., & Inouye, J. (1986). Further specification of the five robust factors of personality. *Journal of personality and social psychology*, *50*(1), 116.

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual review of psychology*, *41*(1), 417-440.

Dryer, D. C. (1999). Getting personal with computers: how to design personalities for agents. *Applied artificial intelligence*, *13*(3), 273-295.

Emojipedia (2019). Retrieved from https://emojipedia.org/

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: a three-factor theory of anthropomorphism. *Psychological review*, *114*(4), 864.

Eysenck, H. J. (1946). The Measurement of Personality [Resume].

Eysenck, H. J., & Himmelweit, H. T. (1946). An experimental study of the reactions of neurotics to experiences of success and failure. *The of general psychology*, *35*(1), 59-75.

Eysenck, H. J. (1991). Dimensions of personality: 16, 5 or 3? Criteria for a taxonomic paradigm. *Personality and individual differences*, *12*(8), 773-790.

Field, D., & Millsap, R. E. (1991). Personality in advanced old age: continuity or change? *Journal of Gerontology*, *46*(6), P299-P308.

Freyd, M. (1924). Introverts and Extroverts. *Psychological Review, 31*(1), 74-87.

Gallaher, P. E. (1992). Individual differences in nonverbal behavior: Dimensions of style. Journal of Personality and Social Psychology, 63, 133–145.

Gifford, R. (1991). Mapping nonverbal behavior on the interpersonal circle. Journal of Personality & Social Psychology, 61, 279–288.

Gill, A. J., & Oberlander, J. (2002). Taking care of the linguistic features of extroversion. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 24, No. 24).

Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. *Review of personality and social psychology*, *2*(1), 141-165.

Goldberg, L. R. (1990). An alternative" description of personality": the big-five factor structure. *Journal of personality and social psychology*, *59*(6), 1216.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological assessment*, *4*(1), 26.

Govers, P. C. M. (2004) Product personality, unpublished doctoral dissertation, Delft University of Technology, Delft, NL

Graaf, M., & Allouch, S. (2014, March). Expectation setting and personality attribution in hri. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction* (pp. 144-145). ACM.

GSMA Intelligence (2019) Retrieved from https://www.gsmaintelligence.com/

Hayes-Roth, B., Sincoff, E., Brownston, L., Huard, R., & Lent, B. (1995, May). Directed improvisation with animated puppets. In *Conference companion on Human factors in computing systems* (pp. 79-80). ACM.

Isbister, K., & Nass, C. (2000). Consistency of personality in interactive characters: verbal cues, non-verbal cues, and user characteristics. *International journal of human-computer studies*, *53*(2), 251-267.

Janlert, L E and Stolterman, E (1997) The character of things, Design Studies Vol 18 pp 297e314

John, O. P. (1990). The "Big Five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 66-100). New York, NY, US: The Guilford Press.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative big five trait taxonomy. *Handbook of personality: Theory and research*, *3*(2), 114-158.

John, O. P., and Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and research*, *2*(1999), 102-138.

Johnson, K. (2019). Facebook Messenger passes 300,000 bots. Retrieved from https://venturebeat.com/2018/05/01/facebook-messenger-passes-300000-bots/

Joosse, M., Lohse, M., Pérez, J. G., & Evers, V. (2013, May). What you do is who you are: The role of task context in perceived social robot personality. In *2013 IEEE International Conference on Robotics and Automation* (pp. 2134-2139). IEEE.

Jordan, P. W. (2002). The personalities of products. *Pleasure with products: beyond usability*, 19-47.

Kammrath, L.K., Ames, D.R., Scholer, A.A. (2007) Keeping up impressions: Inferential rules for impression change across the big five. J. of Experimental Social Psychology 43(3), 450–457.

Karsvall, A. (2002, October). Personality preferences in graphical interface design. In *Proceedings of the second Nordic conference on Human-computer interaction* (pp. 217-218). ACM.

Kestner, J., Leithinger, D., Jung, J., & Petersen, M. (2009, February). Proverbial wallet: tangible interface for financial awareness. In *Proceedings of the 3rd International Conference on Tangible and Embedded Interaction* (pp. 55-56). ACM.

Kiesler, D. J. (1983). The 1982 interpersonal circle: A taxonomy for complementarity in human transactions. *Psychological review*, *90*(3), 185.

Klages, L. (1926). Die Grundlagen der Charakterkunde.

Karampela, M., Tregear, A., & Ansell, J. (2014). When opposites attract: incorporating trait complementarity into the measurement of self-brand personality alignment. In *Association for Consumer Research Conference (ACR 2014)*.

KLM first airline with verified WhatsApp business account. (2019). Retrieved from https://news.klm.com/klm-first-airline-with-verified-whatsapp-business-account/

Lee, K. M., & Nass, C. (2003, April). Designing social presence of social actors in human computer interaction. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 289-296). ACM.

Lee, K. M., Peng, W., Jin, S. A., & Yan, C. (2006). Can robots manifest personality?: An empirical test of personality recognition, social responses, and social presence in human–robot interaction. *Journal of communication*, *56*(4), 754-772.

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of personality and social psychology*, *52*(1), 81.

Mennicken, S., Zihler, O., Juldaschewa, F., Molnar, V., Aggeler, D., & Huang, E. M. (2016, September). It's like living with a friendly stranger: perceptions of personality traits in a smart home. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 120-131). ACM.

Meyer, K. (2016). The Four Dimensions of Tone of Voice. *Nielsen Norman Group. Lu*.

McTear, M., Callejas, Z., & Griol, D. (2016). *The conversational interface: Talking to smart devices*. Springer.

Nass, C., Moon, Y., Fogg, B. J., Reeves, B., & Dryer, D. C. (1995). Can computer personalities be human personalities? *International Journal of Human-Computer Studies*, *43*(2), 223-239.

Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *The Journal of Abnormal and Social Psychology*, *66*(6), 574.

Norman, W. T. (1967). 2800 PERSONALITY TRAIT DESCRIPTORS--NORMATIVE OPERATING CHARACTERISTICS FOR A UNIVERSITY POPULATION.

Norman, D. A. (2004). *Emotional design: Why we love (or hate) everyday things*. Basic Civitas Books.

Novikova, I. A. (2013). Big Five (The Five‐Factor Model and The Five‐Factor Theory). *The Encyclopedia of Cross‐Cultural Psychology*, *1*, 136-138.

Peabody, D., & Goldberg, L. R. (1989). Some determinants of factor structures from personality-trait descriptors. *Journal of personality and social psychology*, *57*(3), 552.

Rutjes, J. A. J. (2013). *Expectations and Human-Robot Interaction. The influence of robot expectations on personality attribution, impressions and anthropomorphism* (Master's thesis, University of Twente).

Saucier, G. (1994). Mini-Markers: A brief version of Goldberg's unipolar Big-Five markers. *Journal of personality assessment*, *63*(3), 506-516.

Saucier, G., & Goldberg, L. R. (1996). The language of personality: Lexical perspectives. *The five-factor model of personality: Theoretical perspectives*, 21-50.

Schmitz, M., Krüger, A., & Schmidt, S. (2007, January). Modelling personality in voices of talking products through prosodic parameters. In *Proceedings of the 12th international conference on Intelligent user interfaces* (pp. 313-316). ACM.

Shmelyov, A. G., & Pokhil'ko, V. I. (1993). A taxonomy-oriented study of Russian personality-trait names. *European Journal of Personality*, *7*(1), 1-17.

Smestad, T. L. (2018). *Personality Matters! Improving The User Experience of Chatbot Interfaces-Personality provides a stable pattern to guide the design and behaviour of conversational agents* (Master's thesis, NTNU).

Somer, O., & Goldberg, L. R. (1999). The structure of Turkish trait-descriptive adjectives. *Journal of personality and social psychology*, *76*(3), 431.

Szarota, P. (1995). Polska Lista Przymiotnikowa (PLP): Narzędzie do diagnozy Pięciu Wielkich czynników osobowości. *Studia Psychologiczne*.

Tapus, A., Ţăpuş, C., & Matarić, M. J. (2008). User—robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy. *Intelligent Service Robotics*, *1*(2), 169.

Tewari, A., Liu, I., Cai, C., & Canny, J. (2012, September). An analysis of the dialogic complexities in designing a question/answering based conversational agent for preschoolers. In *International Conference on Intelligent Virtual Agents* (pp. 36-45). Springer, Berlin, Heidelberg.

Thies, I. M., Menon, N., Magapu, S., Subramony, M., & O'neill, J. (2017, September). How do you want your chatbot? An exploratory Wizard-of-Oz study with young, urban Indians. In *IFIP Conference on Human-Computer Interaction* (pp. 441-459). Springer, Cham.

Togler, J., Hemmert, F., & Wettach, R. (2009, February). Living interfaces: the thrifty faucet. In *Proceedings of the 3rd International Conference on Tangible and Embedded Interaction* (pp. 43-44). ACM.

Trapnell, P. D., & Wiggins, J. S. (1990). Extension of the Interpersonal Adjective Scales to include the Big Five dimensions of personality. *Journal of personality and Social Psychology*, *59*(4), 781.

Tupes, E. C., & Christal, R. E. (1961). Recurrent personality factors based on trait ratings (ASD-TR-61-97). *Lackland Air Force Base, TX: Aeronautical Systems Division, Personnel Laboratory*.

Weizenbaum, J. (1966). ELIZA---a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, *9*(1), 36-45.

Wiggins, J. S. (1995). *Interpersonal Adjective Scales*. Psychological Assessment Resources.

Wilt, J., & Revelle, W. (2008). Extroversion and emotional reactivity. *in press) Handbook of individual differences in social behavior. New York, NY: Guilford Press. ПРИЛОЗИ, 107*.

Yaghoubzadeh, R., Kramer, M., Pitsch, K., & Kopp, S. (2013, August). Virtual agents as daily assistants for elderly or cognitively impaired people. In

International Workshop on Intelligent Virtual Agents (pp. 79-91). Springer, Berlin, Heidelberg.

Yang, K. S., & Bond, M. H. (1990). Exploring implicit personality theories with indigenous or imported constructs: The Chinese case. *Journal of personality and social psychology*, *58*(6), 1087.

Zhou, M. X., Mark, G., Li, J., & Yang, H. (2019). Trusting Virtual Agents: The Effect of Personality. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, *9*(2-3), 10.

Zumstein, D., & Hundertmark, S. (2017). Chatbots—An Interactive Technology For Personalized Communication, Transactions and Services. *IADIS International Journal on WWW/Internet*, *15*(1).

APPENDICES

APPENDIX A: The extroverted chatbot's flow

APPENDIX B: The introverted chatbot's flow

APPENDIX A

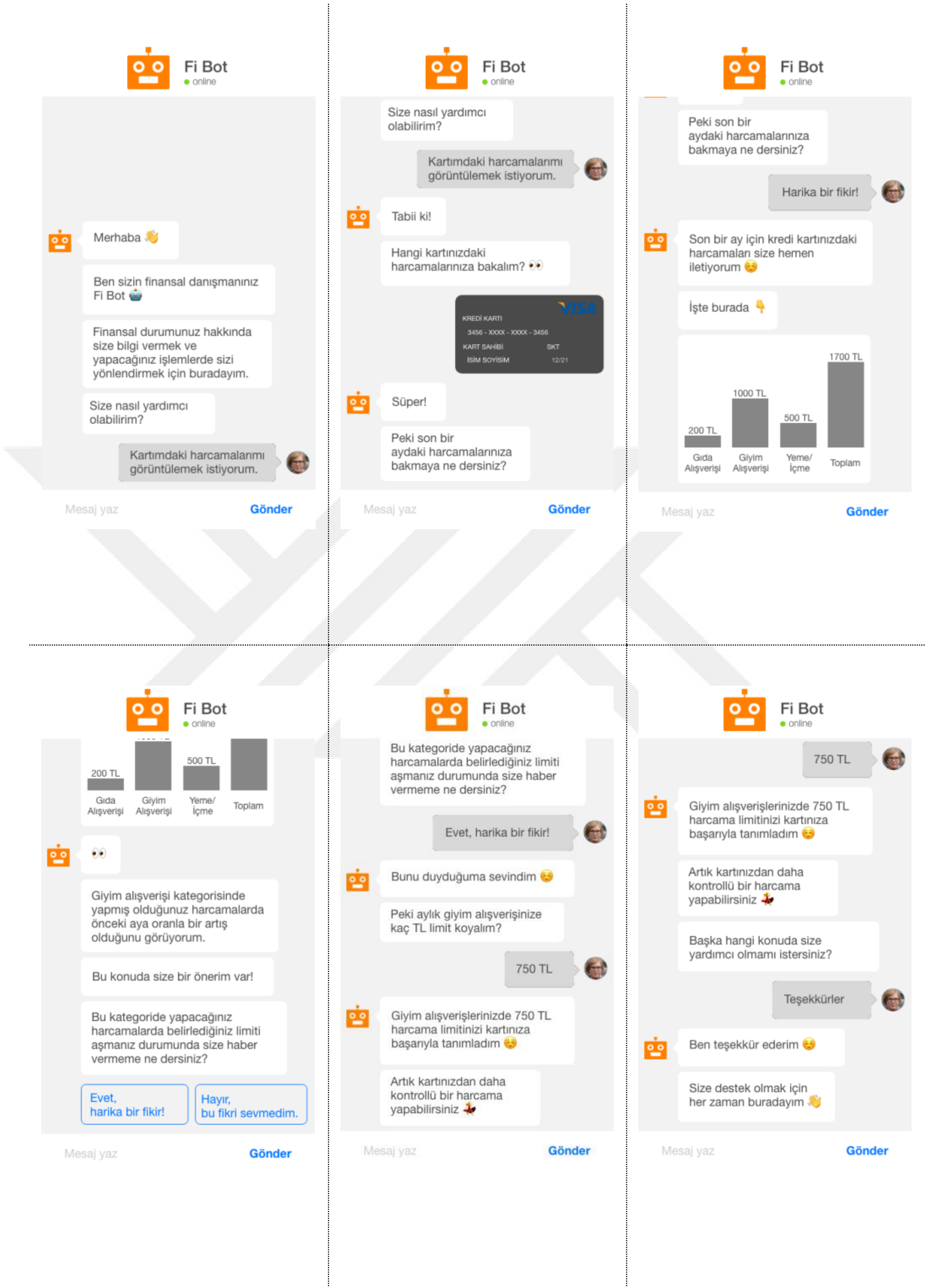

Figure A.1 : The extroverted chatbot’s flow

APPENDIX B

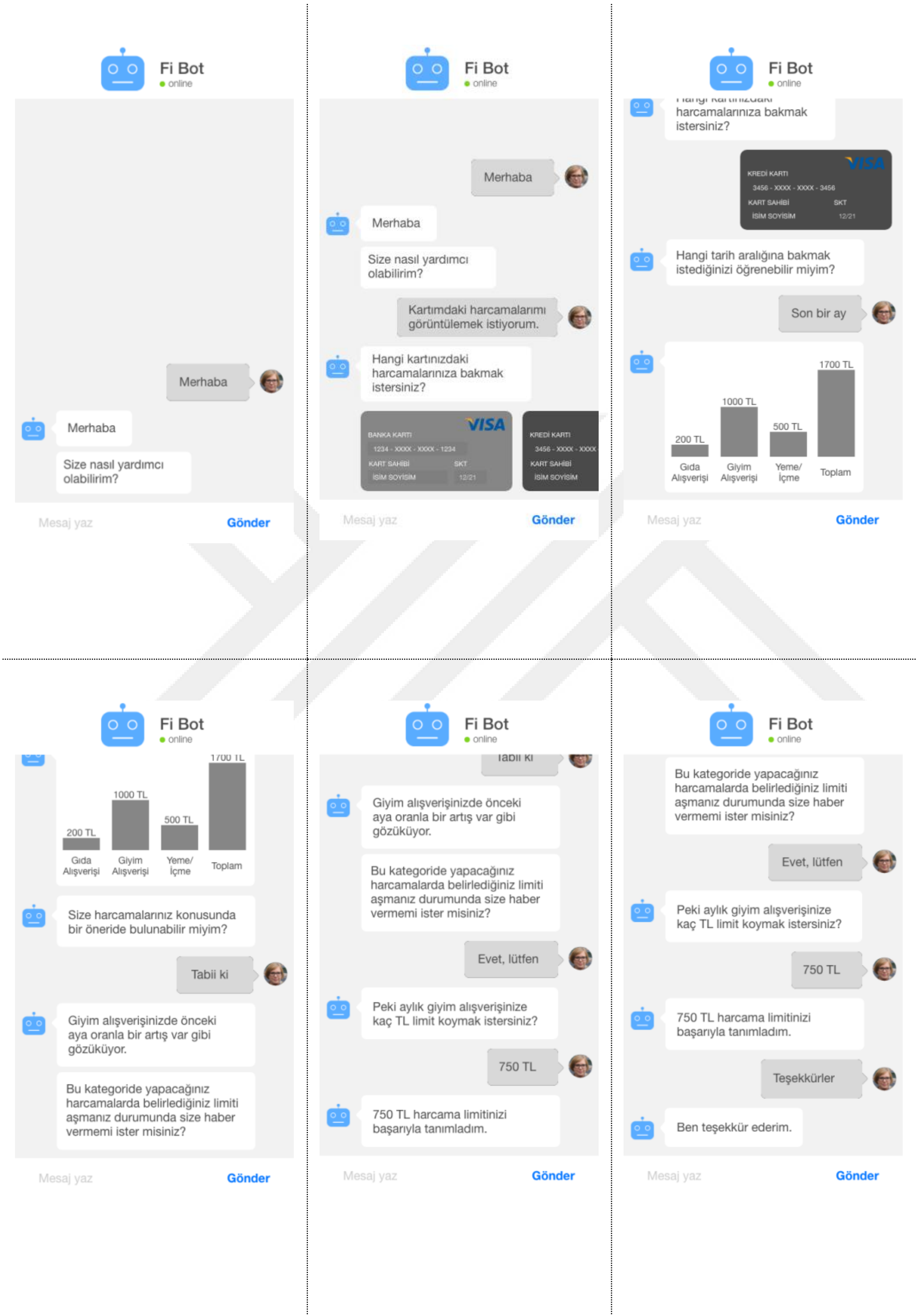

Figure B.1 : The introverted chatbot's flow

CURRICULUM VITAE

Name Surname **:** Gonca İrem Çakıroğlu

Place and Date of Birth **:** Kadıköy, 09.05.1991

E-Mail **:** iremcakiro@gmail.com

EDUCATION **:**

- **B. Sc.** **:** 2016, Istanbul Technical University, Architecture Faculty, Industrial Product Design Department

PROFESSIONAL EXPERIENCE AND REWARDS:

- 2016-2018 Research Assistant at Istanbul Commerce University
- 2018-present User Experience Designer at Garanti BBVA Technology

PUBLICATIONS, PRESENTATIONS AND PATENTS ON THE THESIS:

- **Cakiroglu, I. (2017).** Genders of Products: Creating Genderless Design. *Proceedings of the 9th International Conference Senses and Sensibility*, Funchal
- **Cakiroglu, I., & Pazarbasi, C. K. (2019).** Critical Design in Daily Life: Lifelike Products. *The Design Journal*, *22*(sup1), 1227-1234.